REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
14-12-2005	Final Report	01-Dec-00 - 01-Dec-05

4. TITLE AND SUBTITLE

Mathematical Basis Of Knowledge Discovery And Autonomous Intelligent Architectures - Eye-Tracking And Head-Mounted Display/Tracking Computer System For The Remote Control Of Robots And Manipulators

5a. CONTRACT NUMBER
ISTC Registration No: 1993p

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Professor B.V.Sokolov, Professor F.M.Kulakov

5d. PROJECT NUMBER

5d. TASK NUMBER

5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

St. Petersburg Institute For Informatics & Automation of the Russian Academy of Sciences
39, 14th Liniya
St. Petersburg 199178
Russia

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
PSC 802 BOX 14
FPO 09499-0014

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
ISTC 00-7031-5

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

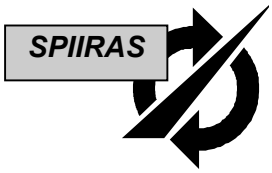
Global awareness (GA) entails the acquisition of data from local to global levels, appropriate fusing of the data, and presentation of that data as useful information. This data will then be fused to fully describe situations of interest such as large transportation systems and complex communication systems. This project specifically aims at developing the mathematical basis, architecture and software techniques implementing particular new technologies to support Global Awareness and comprises six main tasks. Task 5 was:

5. Eye-tracking and head-mounted display/tracking computer system for the remote control of robots and manipulators

15. SUBJECT TERMS

EOARD, Mathematical And Computer Sciences, Computer Programming and Software

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18, NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			PAUL LOSIEWICZ, Ph. D.
UNCLAS	UNCLAS	UNCLAS	UL	78	**19b. TELEPHONE NUMBER** *(Include area code)* +44 20 7514 4474



ST. PETERSBURG INSTITUTE FOR INFORMATICS AND AUTOMATION (SPIIRAS)

EUROPEAN OFFICE OF AEROSPACE RESEARCH AND DEVELOPMENT (EOARD)

Eye-tracking and head-mounted display/tracking computer system for the remote control of robots and manipulators

Final Report

Project # 1992 P

Task 5

Project Manager
Dr.Sci., Professor
B.V.Sokolov

Principal Investigator
Dr.Sci., Professor
F.M.Kulakov

St. Petersburg
November 2003

Contents

Foreword

This Final Report presents the results of activity in framework of Project #1992p, Task 5 "Eye-tracking and head-mounted display/tracking computer system for the remote control of robots and manipulators".

The main objectives of this work are:

- to develop advanced multi-media technology of man-machine interface (MMI) for robot-manipulator telecontrol and other mobile objects control;
- to develop prototypes of systems which track the position/orientation of head (Head Tracking System - HTS), hand (Hand Tracking System - HTS+) and gaze direction (Eye Tracking System - ETS) of man-operator;
- to verify and to validate the developed prototypes and technology with the Hard & Software Complex (HSC) tailored for this aims.

According to the Schedule, in the 11-th and 12-th quarters work was completed for milestone E3 and the results are presented in this Report:

Testing HTS &ETS prototypes and estimation of results (E3-4, 100%).

The results of fabrication and preliminary tests results of HTS & ETS prototypes integrated in the Hard & Software Complex to achieve their appropriate performance while controlling a remote robot-manipulator are given in this Report.

Main results of activity in framework of Project #1992p, Task #5 were presented in oral report and poster reports on 6-th International Seminar “Science and Computing”, 15-17 September 2003, Moscow, Russia.

Introduction

The main objective of task #5 the Project is development of Head Tracking System (HTS) with modification to Hand Tracking System (HTS+) and Eye Tracker System (ETS) for remote robot control.

This work proposes a new intelligent Man-Machine Interface (MMI) using HTS and ETS having the following features:

- Tracking natural man-operator's motions (the head observing motion and the hand controlling motion);
- Realizing for man-operator the work scene imaging with a visual effect close to the holographic one;
- Naturalness of robot teaching by showing man-operator's head and hand motions or manipulated object's motions.

Employing HTS and ETS systems for remote robot control improves solution of the following tasks: the distant robot control, the work scene observation and the teaching of the remote robot control system, see **Chapter 1**.

The results of fabrication of the HTS and HTS+ hardware and software prototypes are presented in **Chapter 2**.

The preliminary test results of HTS and HTS+ prototypes for control of robot-manipulator and robot-like device bearing cameras for generating visual mapping of environment are given in **Chapter 3.**

The results of fabrication of the ETS hardware and software prototypes for robot control are presented in **Chapter 4**.

The preliminary test results of ETS prototypes integrated in the Hardware & Software Complex Facility for testing and control of the 3D Virtual Cursor (VC) and the High Resolution Image Zone (HRIZ) are given in **Chapter 5.**

The experimental data of testing the HTS and HTS+ prototypes are presented in **Appendix 1**.

The experimental data of testing the ETS prototype are presented in **Appendix 2**.

Some applications of results of activity Task #5 and commercial proposals based on developed prototypes HTS, HTS+ and ETS are presented in **Summaries (for Chapter 2-5)** and in **Conclusion**.

Chapter 1. Novel features of MMI with HTS & ETS

Novel features of Man-Machine Interface with Head Tracking System (HTS) and Eye Tracking System (ETS):

- "Presence effect" of man-operator in a remote work zone,
- Master slave manipulator control based on HTS+,
- Supervisory telecontrol based on use a virtual manipulator,
- Supervisory control the target positions of end-of-manipulator tool,
- Biotechnical scanning for generation of environment model,
- Control of high-resolution zone position of displayed image.

1.1. Telecontrol using visual "presence effect" of man in remote work zone

The HTS system can be profitably used in man-machine interface (MMI) of telecontrolled robot-manipulator to provide a highly realistic effect of the man-operator's presence in the robot-manipulator work zone (WZ) actually located at a large distance from the man-operator. This effect has title a "presence effect". It is important to provide a natural interface in both case of biotechnical control and case of supervisory telecontrol when man is supervisor.

The Fig 1.1 presents a block diagram of the proposed man-machine interface based on HTS for master slave robot telecontrol, which has a considerable improvement of "presence effect" of man in a remote WZ.

The stereo TV-cameras with parallel optical axes placed at a distance equal to the baseline between the human eyes produces a stereo-image of the WZ. Data of 6 coordinates from the HTS corresponded to the man-operator's head position/orientation are the input values for double TV-camera control system.

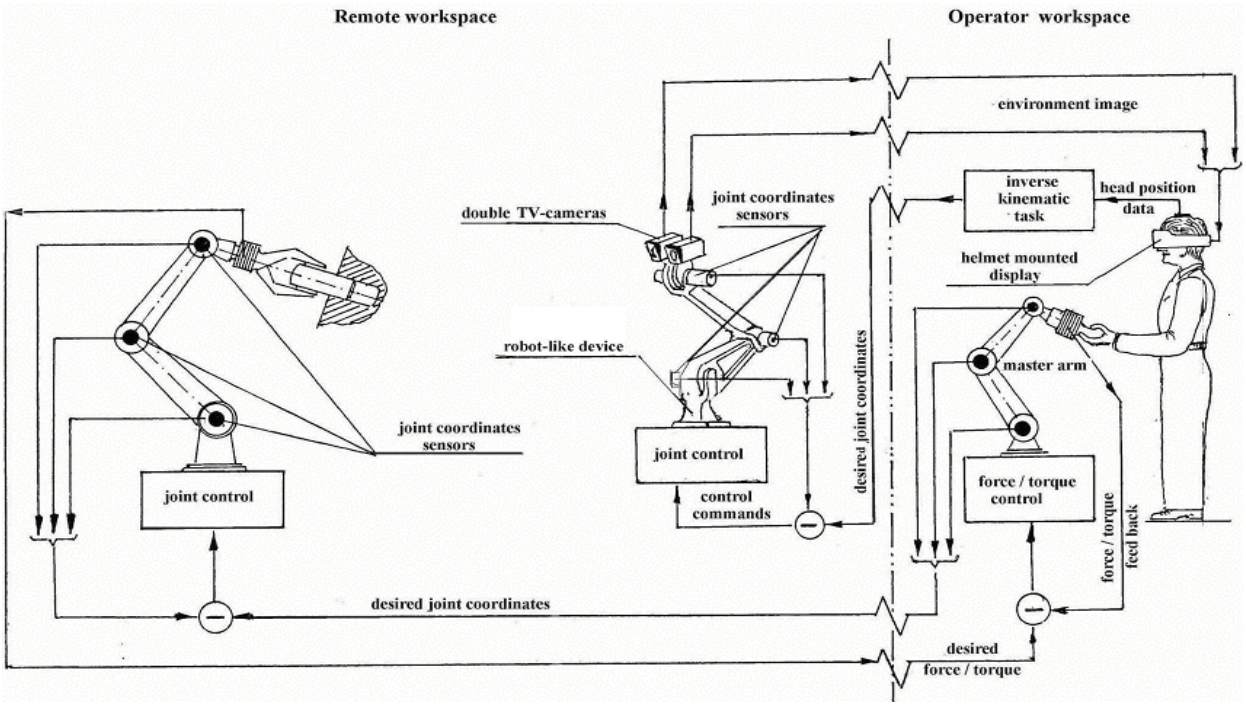


Fig. 1.1 Robot telecontrol with HTS to provide a "precense effect" for man-operator

Tracking of the head position and orientation in real time allows controlling the angle and scale of the 3D image of the remote WZ being viewing through the Helmet-Mounted Display (HMD) generating so-called "pseudo holographic" effect.

This fact can considerably enhance the effect of the operator's direct presence in the remote WZ. Maximal presence effect will be provided the ortostereoscopic conditions are satisfied.

1.2. Master-slave manipulator control based on HTS+

During the experimental studies of telecontrol with HTS a new design of 6D handle for position and speed control was proposed, one based on employment of HTS.

The natural man's hand movement is used for control in this case instead of master-arm. The 6D coordinates for slave arm are generated by Hand Tracking System (HTS+) which tracking hand position/orientation similar to HTS (Fig. 1.2).

Fig. 1.2 Master-slave manipulator control based on HTS+

While using position-speed HTS + 6D handle for control of robot - manipulator one gets a large range and naturalness of movement than those attained with a traditional handle of "Master-Arm" type.

Adequate perception of man natural motions in the environment of the robot control post without any restrictions for his natural (intuitive) behavior maximally realizes his experience, natural reaction and professional skill.

HTS+ provides improvement to Man-Machine Interface in case of the bio-technical master-slave robot telecontrol. Control process will be more natural and simple than in case of traditional master-arm is application.

1.3. Supervisory telecontrol based on virtual robot for prediction & teaching

The modern supervisory telecontrol is realized using so-called virtual manipulator, which is a 3D computer-generated image of real manipulator and it is controlled as real manipulator.

It means that the 3D video image of the virtual manipulator must be "immersed" in a real work zone or, in other term, the real environment video image must be augmented with the virtual manipulator's image (Fig. 1.3).

Man-operator must to percept the virtual manipulator as it is real manipulator being among the WZ objects while he must to control a virtual manipulator.

The problem of augmentation to the real environment the virtual object (manipulator in our case), realized on base of Augmented Reality (AR) technology, was solved by us in the framework of Task 6 of Project 1992p. Fig. 1.4 illustrates the process of augmentation of the virtual manipulator image to that of real environment obtained by the TV cameras.

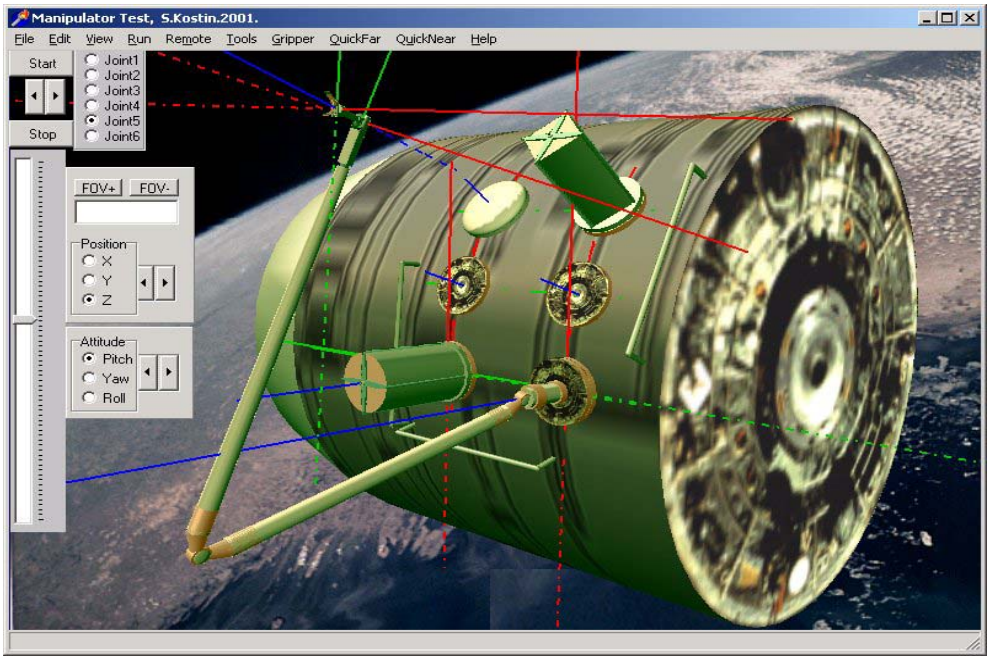


Fig. 1.3 3D-model image of the Orbital Station with a virtual robot image

a). Virtual manipulator

b). Space station mock-up fragment real image with mask

c). Virtual manipulator "immersed" in Space station real image.

Fig. 1.4 Augmented Reality (AR) technology for supervisory robot telecontrol

Employments of the virtual manipulator are the next:

- to plan actions of real manipulator and to check these actions prior to actually execution by real manipulator, i.e. to teach a real manipulator for the task execution;
- to test with virtual manipulator the movement executes by real manipulator, the simultaneously identical control signals are sent to the virtual and real manipulators, that it is necessary in case of large time delay of control signals from control station to real robot, and to stop the real manipulator in time if a control command error takes place.

The more a natural sensing of the virtual manipulator is similar to the real manipulator, the more the supervisory control became natural and simple.

The HTS provides this natural sensing of virtual objects in real time mode. The HTS generates 6 coordinates of man-operator's head position/orientation which are used:

- to set a position/orientation of the virtual observer's head for generation of the corresponding position/orientation (and scale) of the virtual manipulator image and the work zone geometrical models images;

- to set a position/orientation of the mobile stereo-cameras for remote viewing a real work zone image, which must be registries with virtual WZ image.

Head Mounted Display (HMD) presents a virtual manipulator's image immersed in a real WZ image.

The Fig. 1.5 illustrates the proposed Man Machine Interface (MMI) for supervisory telecontrol using the virtual manipulator for verification of a real manipulator actions, for teaching manipulator and, also, for realization of predictive visual feedback in the case of large delay of control signals from the control station to the remote robot-manipulator.

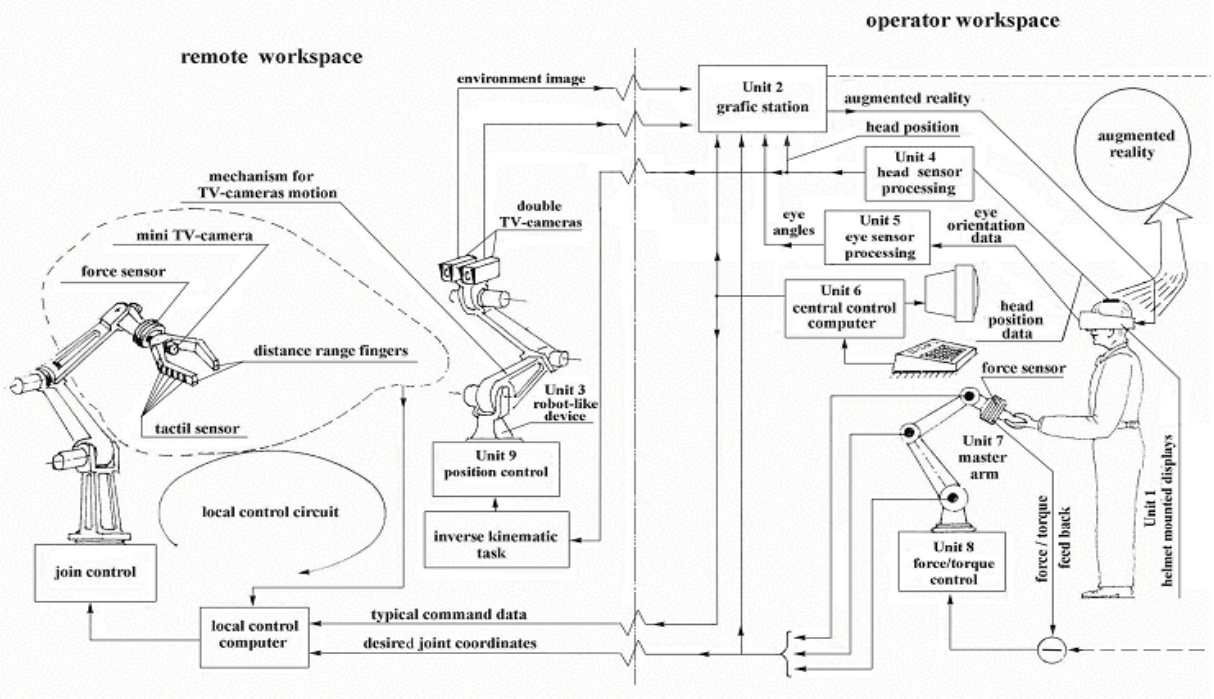


Fig. 1.5 Supervisory Robot Telecontrol and Teaching using HTS

1.4. Based on ETS & HTS the supervisory control of end-of-manipulator tool

The Eye Tracking System (ETS) additional to HTS is a useful means of man-machine interface in case of supervisory telecontrol. They are used for targeting that is for getting the target position of the end-of-manipulator tool and also for getting the coordinates of special points, so-called "points of interest", on surface of obstacles and other environment points of interest.

This data are necessary for the manipulator control system providing an autonomous motion of the manipulator to target position avoiding the obstacles.

The 3D Virtual Cursor (VC) is used for this aim. The VC is a 3D computer synthesized "Virtual pointer". A stereo image of VC is augmented to stereo image of real environment.

The stereo image of real environment is generated with the mobile double TV cameras controlled by HTS.

As a rule, a special 3D joystick is used to control the VC. It generates two position coordinates of the pointer image on displays and the value of the binocular parallax.

We intend to use the ETS measuring man-operator's gaze direction instead of the 3D joystick for VC control. In this case, the turn of a left eyeball is used for control of VC image position on the left display, and the turn of right eyeball is used for control VC image position on the right display (Fig. 1.6).

The TV-images of real environment are outputted on HMD or 3D PC monitor. The computer synthesized images of VC for left and right eyes are outputted also to these displays.

Fig. 1.6 Prototype of ETS for virtual pointer control

If the rendering parameters of VC image are calibrated with the real environment image and the VC stereo image is registered with the environment "point of interest" than the VC images coordinates for each display (left and right) can be used to calculate the depth coordinate and last two coordinates of the point of interest in the double TV cameras coordinate system.

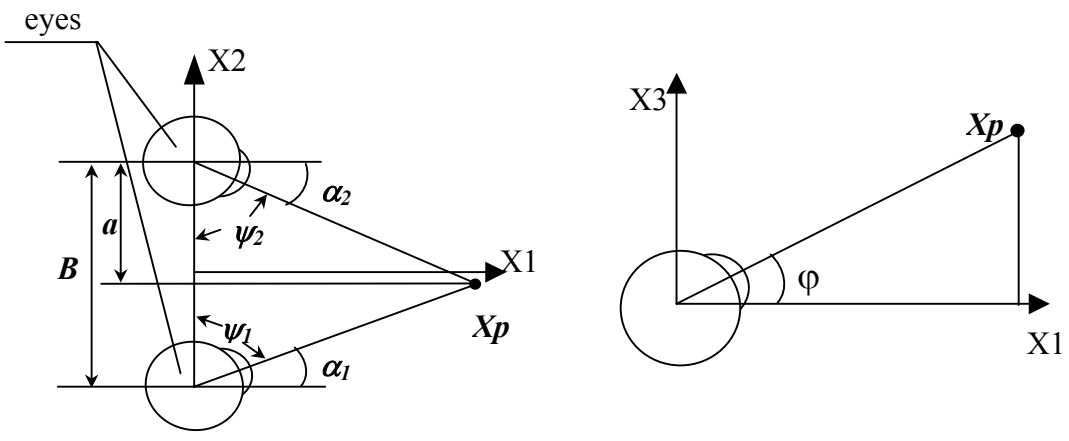
The position of the environment's point of interest in the manipulator coordinate system is calculated by means of data generated by HTS controlling the position/orientation of the mobile double TV cameras in the manipulator's coordinate system.

The possibility of usage an eyeball turn for obtaining position coordinates of the point of interest is based on specific vision functions of human brain. When human views a point of interest the control signals going from the brain to muscles responsible for motion of each eye ball. Then angles of left and right eyeballs are changed so that the optical eye axes intersect at the point of interest.

These eyes angles are measured with ETS, and the position coordinates of VC images are used for calculation of point-of-interest 3D coordinates in the mobile stereo cameras coordinates system.

1.5. Bio-Technical Generation of Environment Geometrical Model

Advanced application of VC control based on ETS is so-called Bio-Technical Generation of Environment Model. This model is a surface dividing the WZ volume of manipulator to two regions: transmitting light rays and opaque ones. Without using such model augmentation of real environment with virtual object (manipulator) is impossible to use the Augmented Reality (AR) technology for supervisory robot telecontrol (Fig. 1.7).



a) vertical plane view b) side view

Fig. 1.7 Biotechnical Eye scanning for generation of Environment geometrical Model

Bio-Technical Generation of the Environment Model is very natural and simple process of WZ visual scanning by eyes or WZ stereo image visual scanning. This process provides generation of a 3D points array on the surface separating the transparent and opaque regions of robot WZ.

It is important to note, that the coordinates of each point of the Environment Model are generated the same way, as those of "point of interest", see Appendix 3.

1.6. ETS for control of High-Resolution Image Zone position of displayed image

A useful application of ETS is VC control for aim to improve resolution of image in a small region near the current gaze position. While operating with a wide size picture a human needs not see all picture with the high-resolution at one moment.

We have experimentally studied a possibility of the High-Resolution Image Zone control (HRIZ control) with ETS (Fig. 1.8).

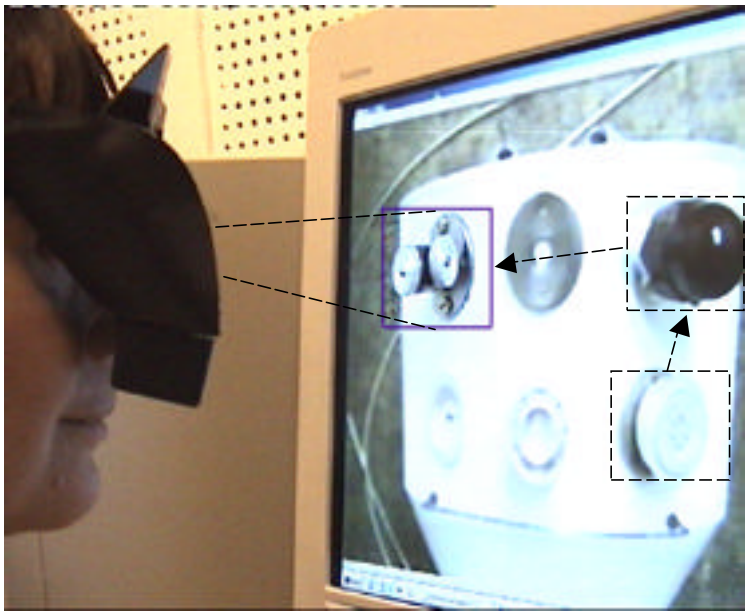


Fig. 1.8 The sequence of High Resolution Image Zone movement by ETS control

This method of dynamic selection enables a considerable reduction of TV image pass-band without spoiling the perceived sharpness of picture. To realize this method one needs equipment enabling permanent tracking of gaze direction. As such equipment the systems HTS and ETS can be used.

It is enough to select a point looked at in a given moment and a small zone around it in which to create image with maximal resolution. But it is important to provide a necessary degree of eye movement synchronism with the movement of the high-resolution image zone (HRIZ). The time lag should not exceed the characteristic eye response time that is about 0,1s.

1.7. Some applications of new MMI system based on ETS & HTS prototypes

Some applications of ETS & HTS for MMI system:

- Advanced computer interface for gesture exchange with PC;
- Pseudo-holographic effect of perception 3D images;
- MMI for the home (office) servicing robot-like devices;
- MMI for telemedicine systems, medical robots, etc;
- Simulators of real time control process (nuclear station, aviation, and others);
- Remote control of camera-head (Web-cameras, security ets);
- MMI for telerobotic control with effect-of-present in remote WZ;
- MMI for multi-robotic systems (control by usage two hands + head).

The Hardware & Software experimental Complex (HSC) for testing developed HTS&ETS prototypes and, also, for testing novel MMI technology for robot telecontrol is showed at Fig. 1.9.

More detailed information on the HTS, HTS+ and ETS prototypes and preliminary test results is presented below in chapters 2-5.

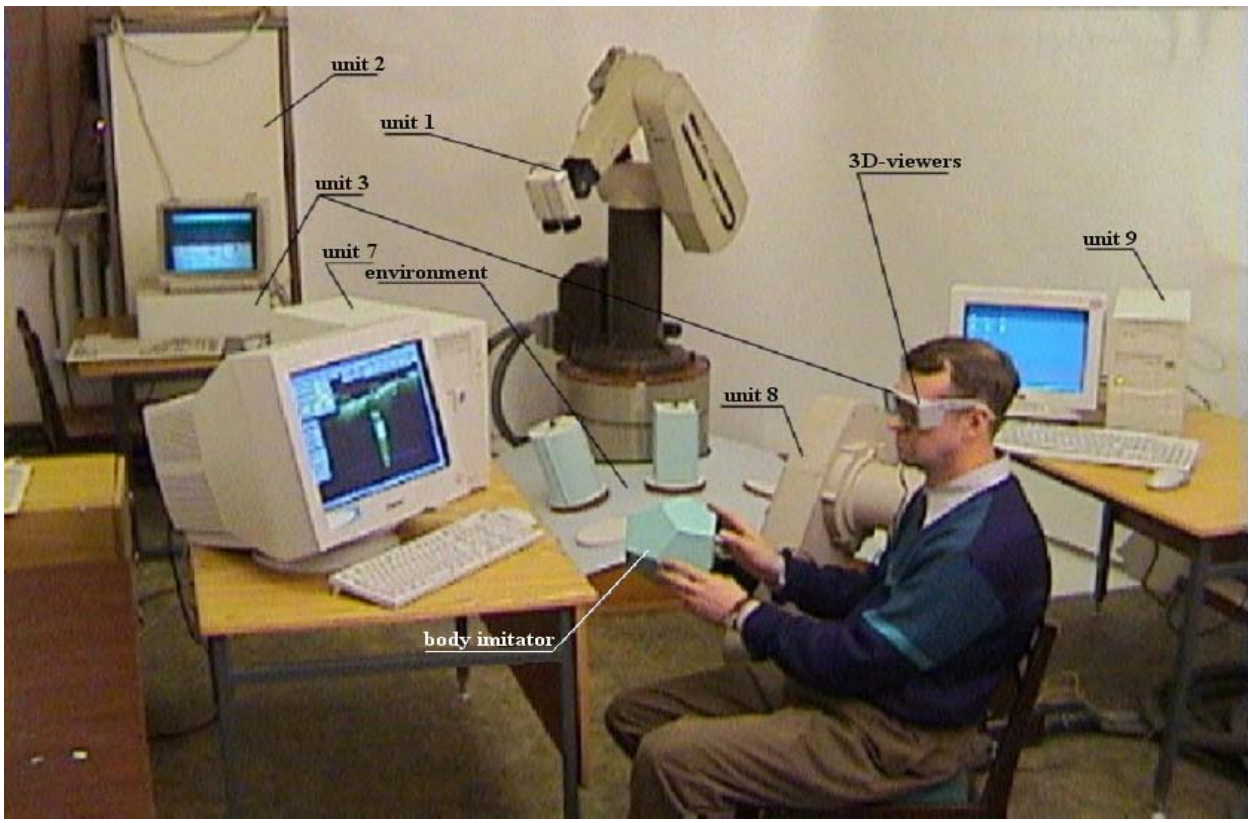


Fig. 1.9 Hardware & Software Complex for testing and verification of MMI based on HTS & ETS

Chapter 2. Hard & software means of man-machine interface for telerobotic using systems tracking man-operator's motion

Preliminary results on a Man-Machine Interface (MMI) development for the robot telecontrol basing on tracking motions of man-operators' head & hand are proposed.

The 3D scene representation with the effect of viewing the pseudo holographic images and the remote robot control by means of natural motions of the man-operator's head and hand have been studied.

Developing the MMI for remote control of robot-manipulator includes solution of the following tasks: viewing of the remote work zone, remote robot control and, also, teaching the robot-manipulator's control system for operating in the remote work zone (WZ).

The main part of this chapter is devoted to the components of the Head Tracking Systems (HTS) and Hand Tracking System (HTS+), their integration into the MMI system and the additional capabilities for man-operator using HTS & HTS+.

Compare with the traditional HTS for aviation purposes, for the robot telecontrol it is essential to enlarge HMB of HTS to provide free movement for operator. Besides that, a comparatively simple and reliable inexpensive HTS operating on PC base is to be developed for robot control.

Three tasks are being solved for MMI realization with HTS and HTS+:

A) Multimedia representation of remote WZ with coordination of image movements with natural motions of head and hand.

B) Adequate perception of man natural motions in the environment of the robot control post without any restrictions for his natural (intuitive) behavior maximally realizing his experience, natural reaction and professional skill. The MMI system also realizes the perception of man handling objects in the real environment of control post.

C) Teaching by showing the MMI system to understand natural motions of man-operator and presentation of the multimedia information in a form maximally comprehensive for a given man. The teaching of a remote robot control system is to be realized in a natural way e.g. showing or demonstrating objects and motions.

2.1. Informative base for the new man-machine interface

There is a great similarity of the Artificial Intellect (AI) system of MMI watching man-operator at his work post to systems for remote watching of robot in real environment. The MMI system for interface with man also uses models (head, hand and manipulated objects) such as traditional systems that recognize the environment objects.

Three information tasks are being solved for MMI realization:

A) Multimedia representation of remote work zone with coordination of 3D image movements with the natural observing motions of man-operator's head realizing for man-operator the WZ imaging with a visual effect close to the holographic one.

B) Adequate perception of man-operator's natural motion in the environment of Control Post (CP) without any restrictions for his natural (intuitive) behavior maximally realizing his experience, natural reaction and professional skill. The ability is also studied of perception of movements, man operating directly at the remote WZ (for example, astronaut in the outer space).

C) Teaching the MMI system in a form maximally comprehensive for a given man-operator.

Teaching a remote robot control system is to be realized in a natural way e.g. showing or demonstrating objects and motions.

There is a great similarity of this MMI system's watching man-operator at CP to a Robot Vision System (RVS) for remote watching real environment. The MMI system deals with a real environment of CP and it must use models both of men and objects in CP. The methods used for structural representation of multimedia data in this MMI system are analogous to the modern methods for structural data compression of images (MPEG-4, 7).

The informative base for MMI data is the 3D models of the following objects:
- head, face, eyes, hands and other parts of man-operator's body;
- spatial shapes of motions executed by head or hands of man;
- reference devices on head and hand of man-operator or the manipulation objects (tools) that a man operates with at CP;
- handled objects at work zone of astronaut (or mock-ups of real objects).

2.1.1. The frame-structural model of objects in MMI

In the HTS image processing and computation of head pose coordinates (position & orientations) are made basing on a priori 3D wire-frame head model.

As the model of head, face, hand and the Reference Device Unit (RDU) on head or in operator's hand we propose 3D graph-like structure which vertices are tables of parameters (or frames) describing properties of each artificial or actual reference mark (specific feature) [1, 2].

This Frame-Structural Model (FSM) stores simultaneously two kinds of information:

1). Data on characteristic properties of mark images needed for automatic selection and identification of images;

2). Parameters defining configuration of marks' mutual positions in a real object or head or hand specific features.

Therefore, the basic properties of FSM are analogous to both types of known descriptions: visual graphs [3] and frame descriptions [4]. Some analogies are models of crystalline structures and models of molecules wherein configuration of links and type of atoms in the nodes define properties of substance.

For example, FSM model configuration is described by a set of relative spacings. Spacing between i, j reference marks in the model (RM_{ij}) are normalized relative to the basic spacing (RM_b) between the marks:

$$RM^{n}{}_{ij} = \frac{RM_{ij}}{RM_b}$$

Where: RM_b – basic spacing length equal, e.g. to the maximal spacing (RM_{ij}) or spacing between specific marks in object.

Besides, configuration is described by a set of spatial angles formed by wire ribs connecting the nearest (neighbor) marks, between radii from the k^{th} mark to i, j marks ($\alpha M^{k}{}_{ij}$).

The parameters describing properties of a mark image and object configuration that are needed for identification and determination of coordinates of mark images in the camera image plane are stored in a frame (M. Minsky)[5].

For example, the first level of the frame description of any object feature as a result of 3D model teaching consists of the following parameters (Fig. 2.1):

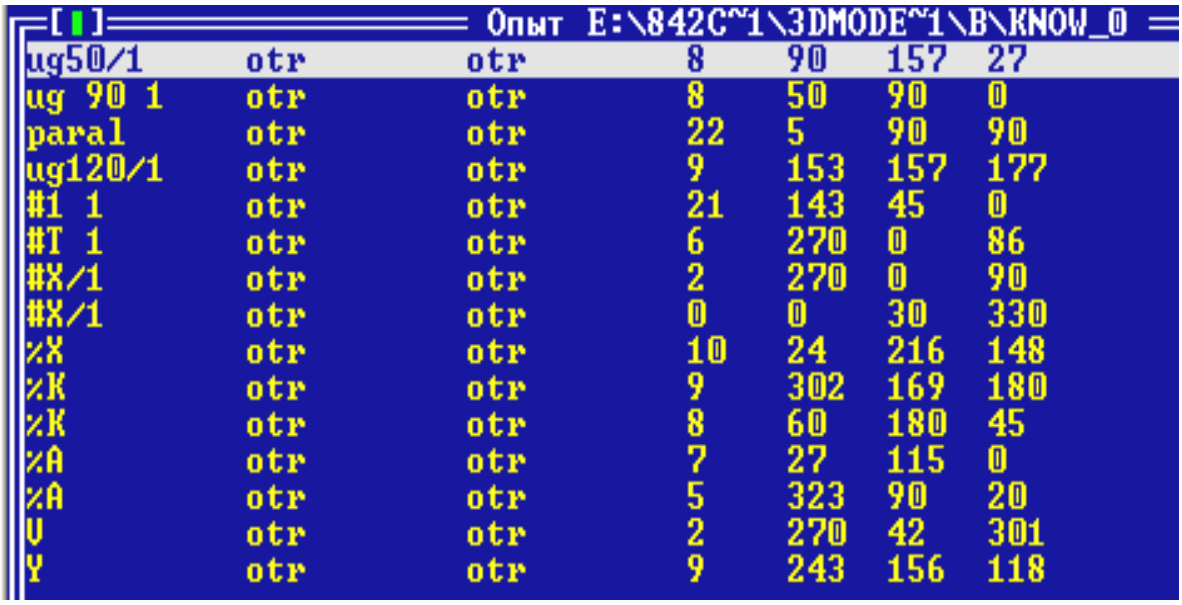


Fig. 2.1 An example of 3D wire-model description (level 1)

A). Mark image feature parameters:

A1. NUMBER - index for feature identification (Ug90, Shape),
A2. TYPE - name of feature (angle =”ug 90”, shape T = “#T” etc.),
A3. NAME - name of a priory elements (line segment = “otr”),
A4. SIZE - dimensions of feature (8 … 22),
A5. ANGLE - angle of element rotation (50….270),
A6. ORIENTATION - orientation of a priory elements (0….360),
A7. COLORS - R, G, B feature color,
A8. SHAPE - feature gradient function.

B). Reference device configuration parameters:

B1. LINK - length of spacing between the i^{th} mark and neighbor j^{th} mark (RM^{n}_{ij});
B2. ANGLE - between links from the k^{th} mark to i^{th} and j^{th} marks (αM^{k}_{ij});
B3. SCALE - proportion of the model to the real object (SM).

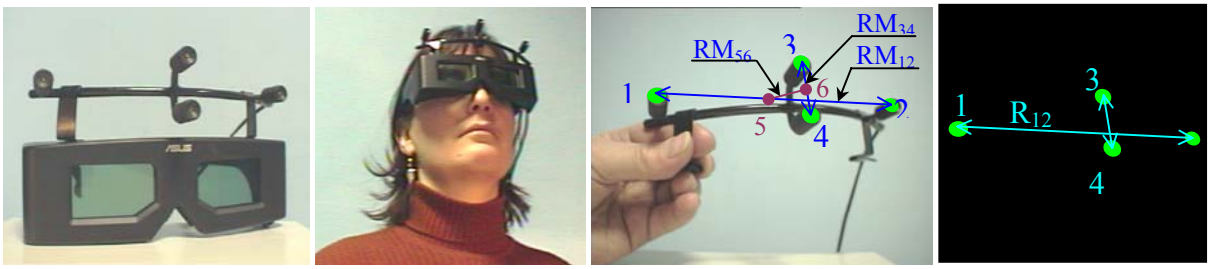
C). Parameters describing configuration of relative positions of reference device and camera:

C1. POSITION - current translation vector (Toc) of the model’s centre to the origin of the camera coordinate system (Xc,Yc,Zc).
C2. DISTANCE - current distance (Doc) between model’s centre and the origin of the camera coordinate system.
C4. ORIENT - current rotation matrix (Roc) storing angular coordinates of the model in the camera coordinate system.

Five levels of 3D model description are used in the HTS prototype:
- level 1 – a priory level of simplest elements (line segments) connected in angles and joints;
- level 2 – the level of description simple shape of wire-frame figures (triangles, brackets, letters);
- levels 3, 4 – the middle levels of wire-frame shape description;
- level 5 – the upper level of description 3D model of RDU as a whole.

The main advantage of the proposed 3D FSM is a relatively simple algorithm for generation of 2D models of object’s images – as projections of 3D wire model upon the camera image plane. For that, in the process of the algorithm’s operation, parameters of 3D model are supplied with parameters of relative positions of 3D model and camera unit (CU) model.

For example, a simple 3D wire-model of RDU on of man-operator’s head is presented on Fig. 2.2.



a) RDU on the stereo viewers, b) RDU with stereo viewers on operator's head, c) Photo of RDU and its 3D model, d) 2D model of RDU

Fig. 2.2 Active RDU and its 3D and 2D models

The geometric 3D wire-frame model is built and displayed by computer means of 3D graphics (Direct-X, VRML or OpenGL) marking each graph vertex with individual number. The 3D model configuration, defined by relative spacing and spatial angles between links, corresponds, with a specified accuracy, to the 3D configuration of relative positions of characteristic features on real object (RDU or head). The examples of more complicated 3D model are given in Fig. 2.3.

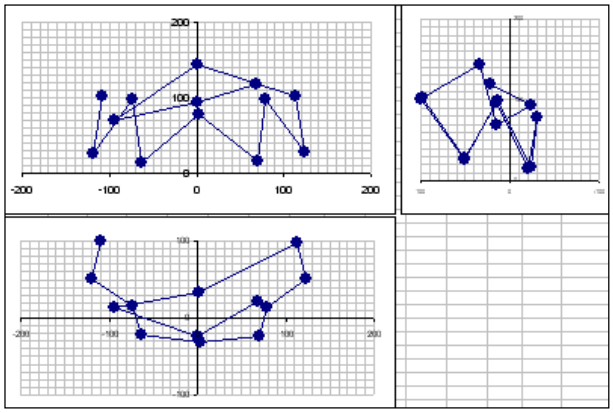
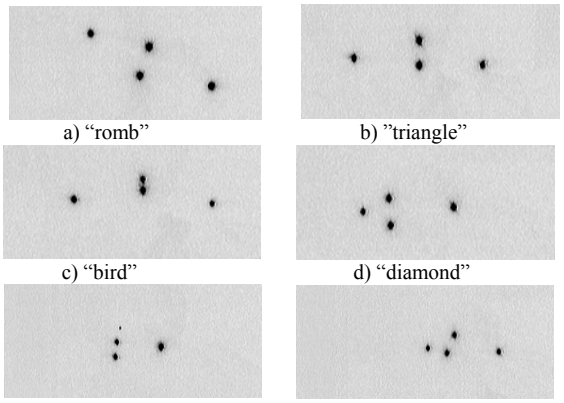


Fig. 2.3 The HTS prototype, 3D RDU model (10 referent points)

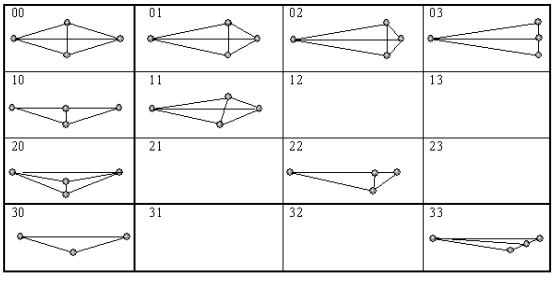
The interpretation and teaching the MMI system with a frame-structural description in a language understandable to operator become simpler. The teaching process may be realized in different ways, as one – by simple showing the objects to MMI system.

The set of 2D models of 4-mark RDU may be generated from the statistics of experimental images from HTS, see Fig. 2.4 (a-f). The type (name) of each 2D RDU model is suggested by man-operator in teaching process for creating 2D RDU model [6, 7]. This number of 2D models (set of 2D models) describes all main angular and linear relative positions of RDU and camera unit (CU) met with in the process of HTS work.

Further, these images will be referred to as 2D RDU models, see Fig 2.4 (j).



a) "romb" b) "triangle" c) "bird" d) "diamond" e) "aeroplane" f) "mouse"



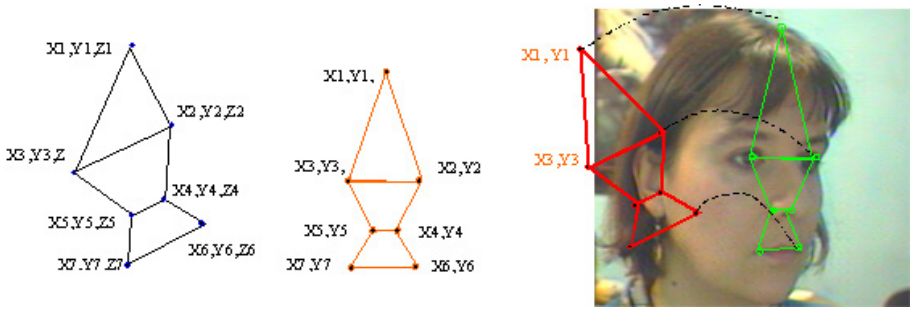
00 – initial RDU pose fronting on CU;
00 - 03 – poses of RDU azimuth rotation (to left);
10 - 30 – poses of RDU rotation for elevation (down wards);
11 - 33 – poses of RDU diagonal rotation (left and down).

j) Set of 2D models RDU (10 main RDU poses relative CU)

Fig. 2.4 Some images of 4-mark RDU in typical poses (phases)

Each 2D RDU model is assigned to a stationary, as we call it, configuration (phase or pose) of relative positions of RDU and CU, one in which RDU image pattern does not change significantly. To each phase (pose) there corresponds its own range of spatial angles giving RDU (head) orientation respecting the camera. With in that range we use only parametric variations without change of any quality new 2D RDU model. Using the 3D model, e.g. for identification of reference mark images, at present is accomplished in two stages.

For reducing mass of computation the image processing is performed in local zones determined with 2D model, see Fig. 2.5. Basing on knowledge of precise measurements of mentioned the head model parameters (head spatial position and orientation) are computed.



a) 3D man-operator' head model b) 2D head model in initial position c) comparison of real image with 2D model of head

Fig. 2.5 Computation of head spatial position by comparison of image with the 2D head model

Comparing images with the 3D model may be realized in the following ways:
- correlating 3D model's spatial coordinates with computed image-based 3D coordinates (characteristic points of real images);
- correlating 2D model projection on image plane with 2D characteristic points of real images.

For deriving estimate (S) of difference in positions of marks (specific features) on the real image [$f(x,y)$] and on the 2D model image [$g(x,y)$], generated by HTS software module, the following expression is used:

$$s = \sum_u \sum_v f(u,v) \oplus g(u,v)$$

To make simpler expression for estimating of coincidence of model and real images the equation is binarized to give:

$$f(u,v) \oplus g(u,v) = \begin{cases} 0, for\ f(u,v) = g(u,v) \\ const., for\ f(u,v) \neq g(u,v) \end{cases}$$

The less value of S the more correct the model and real images of RDU or head. This estimation depends from area of image occupied by RDU image within a video frame [1].

2.2. The HTS prototype design for telerobotic MMI

For the robot telecontrol it is essential to enlarge HMB of HTS to provide free movement for operator. Besides that, a comparatively simple and reliable inexpensive HTS operating on PC base is to be developed for robot control.

2.2.1. *Operational principle of HTS hardware prototype*

The HTS hardware is represented by a functional scheme combining two varieties of optical HTS: active and passive. Consider the basic principles of active HTS and differences with passive one (Fig. 2.6).

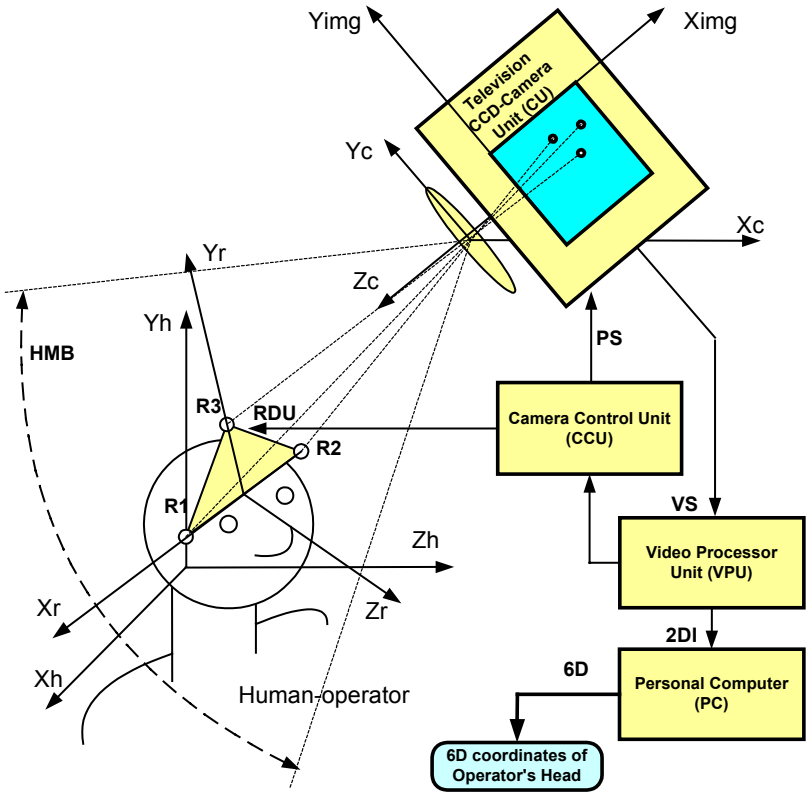


Fig. 2.6 Functional diagram of HTS prototype

1). Human operator performs natural head movements in the Head Motion Box (HMB) volume. In the same time, a helmet-mounted reference device unit (RDU) moves in the HMB for 6 coordinates: three linear translations (x_h, y_h, z_h) and three rotation turn (φ_{xh}, φ_{yh}, φ_{zh}) in the head coordinate system (*Xh*, *Yh*, *Zh*).

2). RDU module has 3 reference marks R1-R3 (IR LEDs for active variety of HTS and color for passive one) rigidly mounted on the RDU base and coordinates of each reference mark (xr, yr, zr) are exactly known in the RDU system of coordinates (Xr, Yr, Zr).

3). CCD-Camera Unit (CU), rigidly mounted on the control console base or PC monitor, is aligned so that the reference marks of RDU always remain in the camera FOV while head of operator moves within the HMB.

4). Reference mark images projected on camera's Focal Plane Array (FPA) will have coordinates (*Ximg*, *Yimg*) in the image (camera) coordinate system (*Xc*, *Yc*).

5). The CU control, power supply and synchronize are executed by the camera control unit (CCU). For the active HTS the most important CCU function is synchronization of camera exposition with pulsed emission of IR LEDs. That makes possible a considerable shortening of exposition time (to 5 μs and less) resulting in rejection factor about 1000 against background interference.

6). Camera video signals come for digital processing to the Video Processor Unit (VPU), implemented as several standard PCI cards at PC Pentium-3 (4). Basic VPU functions are the following: video signal digitization, filtering and selection of reference images on the background and, also, calculating center coordinates with sub-pixel accuracy.

7). Reference marks' coordinates (*Ximg*, *Yimg*) from VPU are entered the PC memory. For known internal and external parameters of the cameras optical system, which are corrected in the procedure of camera calibration, and for coordinates of reference mark images (*Ximg*, *Yimg*) 3D coordinates of the real position of reference mark in the camera coordinate system (*Xc*, *Yc*, *Zc*) are computed.

8). Using the HTS prototype software, installed at the PC, the reference mark images are processed for selection and identification, and RDU position and orientation in the CU coordinate system are computed too.

2.2.2. *Versions of HTS hardware design*

The design of RDU mounted on the miniature telephone garniture on head of human operator controlling a remote robot-manipulator (Fig. 2.7a) was demonstrated at International Conferences on March 2002 [3].

a). The design of RDU mounted on the telephone garniture.

b). Passive RDU design with the stereo viewers

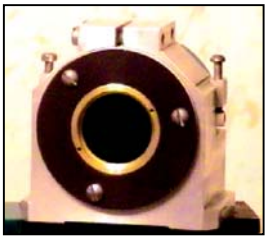
c). The passive RDU design of head ring (band) enabling azimuth head turn of 360 degrees.

Fig. 2.7 Design of RDU for HTS prototype

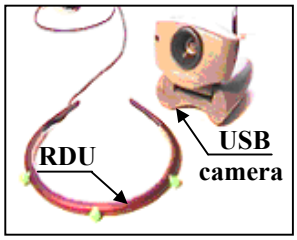
This RDU design is universal one operating both in active and passive HTS modes. Some original solutions were used in the design.
An RDU design combined with the stereo viewers is shown in Fig. 2.7 (b). The RDU design was developed in the form of head ring (band) enabling acquisition of azimuthal head turn in the range of 360 degrees. Combination of color passive marks is identified in this case Fig. 2.7 (c).

The prototype of the HTS camera unit (CU) has a modular design enabling utilization of 1 or 2 cameras, black-and-white or color, what makes possible testing different algorithms of HTS operation (Fig. 2.8).

a). B/w camera of HTS prototype



b). High frequency HTS camera (100Hz)



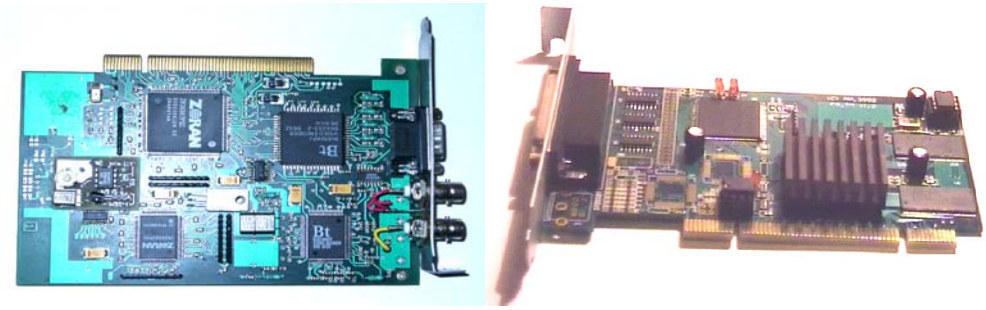
c). HTS's USB-camera with light weight RDU

Fig. 2.8 Design of Camera Units of HTS prototype

Both for active and passive varieties of the HTS prototype we use miniature commercial video cameras: black & white 768x576 pix. and color one (PAL) with resolution no worse than 400 TV lines (Fig. 2.8 c).

The additional study was carried out for establishing a possibility of creating a camera with the speed four times higher than conventional ones, with the frame rate 100 Hz (Fig. 2.8 b) [6].

The video processor's (VPU) hardware is implemented on standard cards in PCI slots (see design of two VPU cards in Fig. 2.9).



a). The input/output VPU card b). The processing VPU card

Fig. 2.9 The designs of the VPU cards

2.3. Description of the HTS prototype's algorithm and software

The significant features of HTS prototype algorithm are the following:

1). A 3D frame-structured model (FSM) of reference device for active and passive HTS types (for the markless HTS – model of operator's face / head) generating it basing on 2D models of images in the camera system of coordinates. Using a 3D model increases reliability of identification of reference marks (characteristic features of face) on the real background.

2). A prediction algorithm for obtaining the most probable places of reference marks on the camera image plane basing on determined speed vectors of their movement.

3). Using color gradient selection of passive reference marks for their localization and identification on the background and, also, for obtaining coordinates of reference mark image centroids with subpixel accuracy.

2.3.1. Description of structural scheme of the HTS prototype's algorithm

A). On the first stage of operation the algorithm performs a spatial filtering (localizing) of mark image zones [8]. Besides, the filtering is performed of the mark's (head's) moving images against stationary background using temporal (template) filter [6]. While using the tracking procedure the localization of mark zones is performed predicting a future image position.

B). After localization of zones of a possible mark image position the mark selection is performed for the beforehand known features: gradient of brightness, colour gradient, size, orientation and shapes of image elements (lines, arcs) and, also for texture [9].

C). On the next stage of the algorithm's operation the identification is performed of each selected mark image – attributing them their numbers. As in the selection the beforehand known parameters of marks are used for that. Besides, the information is used on mutual spatial positions of marks (head specific features for a case of the markless HTS+) that are stored in RDU (head or face) 3D model [6].

This 3D model serves for filtering, selection and identification of mark (head) images at the background of various interference possible in real conditions. Moreover, only those zones of a frame are subject to analysis in which mark images are expected.

D). Then mark image centre coordinates are computed with subpixel accuracy for input images of marks (specific features). Now, a selective secondary filtering is performed at those zones in which the identified marks are. This significantly increases the algorithm's speed and robustness against interference.

E). On the next stage of the algorithm's operation a preliminary estimation of head position and orientation (pose) respecting CU is performed, the fast linear computation is used aided by the RDU 3D model.

This preliminary head pose estimation reduces probability of significant mistakes by image analysis and helps to eliminate false images alike to the marks but not satisfying the model compositionally.

Besides, with excessive number of reference marks (specific features) the preliminary pose estimation enables taking for computation those marks that provide the maximal accuracy of computing spatial coordinates.

F). An accurate computation of head spatial coordinates in the CU system is accomplished by different methods depending on number of cameras in CU. When two and more cameras are used a widely known triangulation calculation method is used, one variant of which is given in [10]. In a case of more economical version of single-camera CU a system of nonlinear equation is solved by an iterative method described earlier in [11].

G). For increasing speed and robustness against interference the prediction of future mark positions is introduced for what mark speed vectors are calculated basing on computed coordinates. Speed measurement results are used for predicting positions of local zones containing reference marks in which image processing is performed not losing time for processing the whole frame. Besides, an adaptive digital predicting filter (Kalman filter or its modifications) makes possible reducing dynamic measurement errors.

H). The HTS algorithm has an additional potential for analysis of head gestures and face mimic information which, in our case, may be used for exchange with other operators and PC (controlling PC by gestures).

2.3.2. Some features in displaying 3D images using a system tracking operator's head movements

Among known systems for displaying 3D images the most widespread are those that use stereo viewers. In these systems a stereo-couple images are presented alternatively on the monitor screen 1 by commutating synchronously light-switch goggles 2 (see Fig. 2.10).

These images, produced to left and right eyes (at the rate no less than 80-100 Hz), are perceived by operator as a single 3D image 3, (triangle ACD) seen at some distance from the screen. Consider then the effect of operator's displacement on the perceived 3D image (taking that the operator work post is equipped with an HTS) giving angular and linear coordinates of rotation centers for left and right eyes of operator in the monitor screen coordinate system.

Assume that operator's eyes shift at H on condition that the centers of rotation O_{L1} and O_{R1} remain in plane 4.

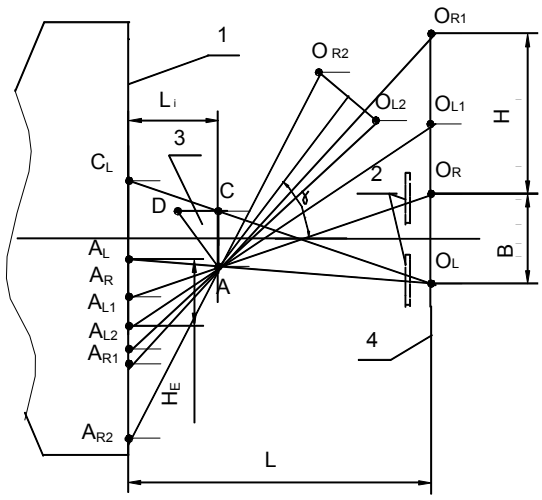


Fig. 2.10 Diagram of 3D image generation on the monitor screen

For the 3D image's remaining in the same place at such an eye shift the stereo-couple images on the monitor screen should be shifted in the opposite direction at $A_R A_{R1} = H_E$ given by expression:

$$H_E/H = L_I/(L_0-L_I) \quad (2.1)$$

From (2.1) we get:

$$H_E = H \cdot \Delta/B$$

While dealing with 3D images the operator is naturally look at them from different sides.
The principle of this method consists in the following:
- man-operator observes a steady 3D image from the middle position;
- as soon as the head begin movement relative to the 3D image HTS outputs change in the angle of sight γ (between normal to the eye base and the screen) and controls the image to turn oppositely to the head at an angle θ related to γ by equation:

$$\theta = \phi(\gamma)$$

Specifically, θ and γ may be proportional with a constant factor K:

$$\theta = K\gamma$$

The full field of view Ω for 3D image will be:

$$\Omega = \gamma + \theta = (K + 1)\,\gamma$$

Assume e.g. K = 2. Then for angles of sight γ = ±30° 3D image will turn at θ = ±60° and the operator will be actually able to survey 3D image from the three sides (Ω = ±90°). For surveying 3D image from all sides (Ω = ±180°) the operator is to turn from the normal to the screen at γ = ±60°. It is, naturally, possible for HTS having a range (yaw) no less than ±60°. The effect of a pseudo-holographic image is achieved not so much with photographic fidelity but providing coordination of movements of head and the stereo image moving under control of HTS (Fig 2.11), see for more detail [12, 3].

Fig. 2.11 The holographic-like effect of 3D image viewing with HTS prototype

Moreover, some degree of conventionality not only does not obstruct the plausibility of picture but also evokes imagination of man. That is especially well seen with dynamic scene effects. Experiments with HTS & ETS prototype showed that graphic controllers of types V3800 Ultra Deluxe and ASUS V6800 GeForce used in PC provide images in a format of 1024x1024 pixels.

A tolerable quality of stereo images may be achieved with frame rate no less than 100 Hz. The format of images produced with such a rate should not be worse than 1024x1024 pixels.

The experiments showed that estimating thresholds of stereopsis with high accuracy requires displays with the screen no less than 19”.

It is known that the stereo effect increases while observing mobile objects. Therefore, it is expedient to use moving test objects, withdrawing and approaching observer, for appraising the boundary of disparity.

2.3.3. Main structure of HTS prototype's software

The HTS' main software for image processing and parameters adjustment with 3D model data in real time mode is presented in Fig. 2.12.

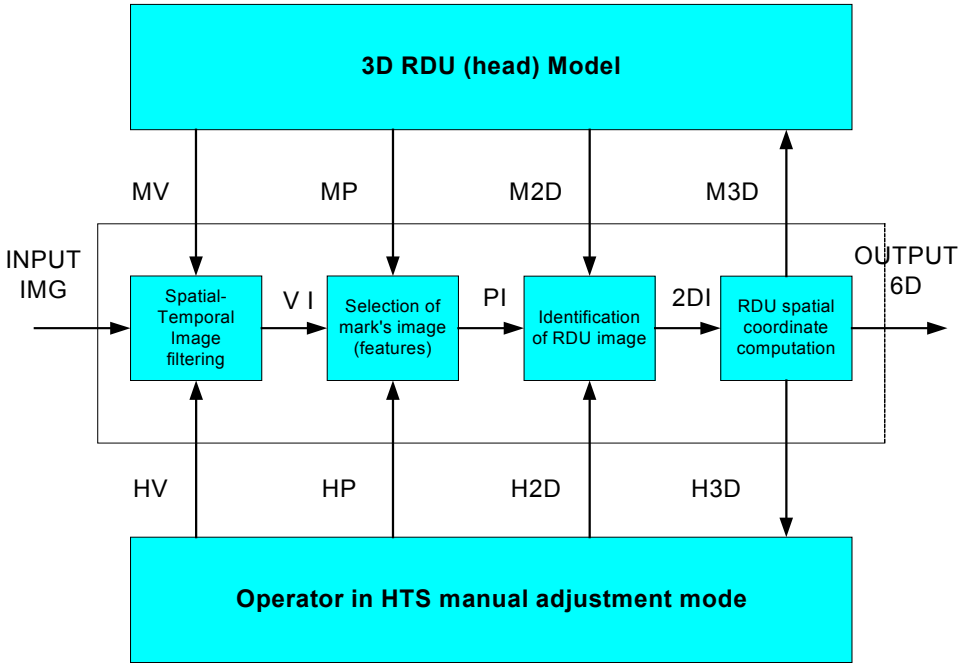


Fig. 2.12 Functional diagram of HTS algorithm

1). Basic information data:
- Input mark images (INPUT IMG);
- Image after spatial-temporal filtering (VI);
- Image after mark feature image selection (PI);
- Image after reference mark identification (2DI);
- Output coordinates of RDU (head) position and orientation (OUTPUT 6D).

2). Auxiliary 3D model data for automatic algorithm's adjustment in real time mode:
- Adjustment parameters for image-temporal filter with image movement prediction (MV);
- Adjustment parameters for variation of mark image properties (size, color, shape, orientation etc.) for RDU (head) image (MP);
- Data for comparison of mark images with 2D model image (M2D);
- Data for correcting 3D RDU model from results of computation (M3D).

3). Auxiliary man-operator's data for manual algorithm's adjustment:
- Adjustment parameters for spatial-temporal filtering (HV);
- Adjustment parameters for specific feature selection (HP);
- Mark image coordinates identified by operator (H2D);
- Displaying results of spatial coordinates' computation for operator (H3D).

The main structure of software for control of robot-manipulators (RC SW) including more than 48 SW modules is presented in [4].

Examples of initial image (IMG_INPT) or filtered image (IMG_FILT) displayed during the program debugging see in Fig. 2.13.

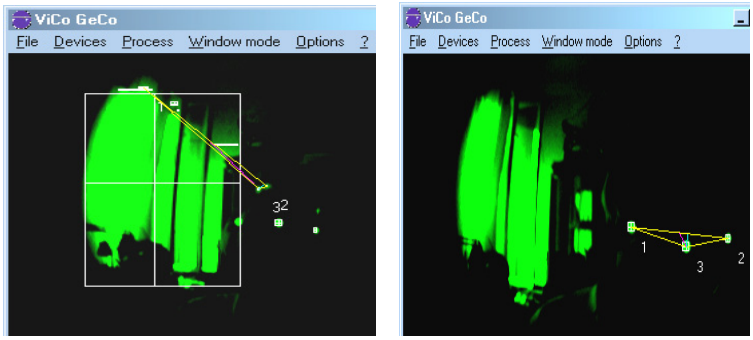
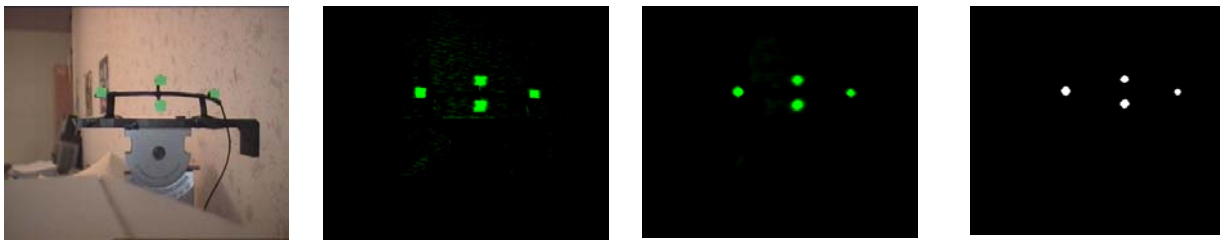


Fig. 2.13 An image of sunlit window and 3-mark RDU before and after filtering

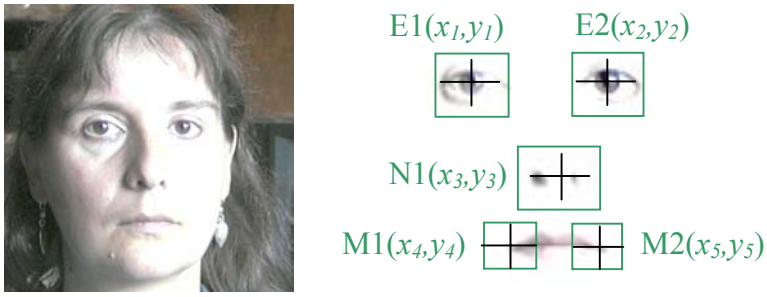
The colour information is very useful for selecting reference marks on a complex background, the more that when bright light sources are present (lamps, sun reflections) in cameras' FOV. The method using colour selection works fast but needs more reliability in conditions of unknown a priori illuminance pattern (intensity and spectral characteristics), see Fig. 2.14.



a). initial image b). selection of colour marks c). median filtering d). binarization of image

Fig. 2.14 Color selection in unknown a priori illumination conditions

For the perspective, markless, HTS the search is accomplished with respect to additional information on specific features of operator's face (head). In this case, a need arises to employ a 3D head (face) model for spatial filtering of image (see Fig. 2.15).



a) initial image b) selection of feature zones

Fig. 2.15 Selection of features in head images by markless HTS prototype SW

Basic technical specifications of the HTS prototype:

1. Measured data (6D coordinates) - 3 angles and 3 linear movements.
2. Work range for angles, no less - 180°, 75°, 60° (yaw, pitch, roll).
3. RMS angular, error no more - 10 angle min,
4. Size of work zone (HMB), no less - 400×400×400 mm.
5. RMS linear, error no more - 0,1 mm.
6. Measurement data are transmitted to external devices via RS-232 serial interface.
7. Data rate from 25 Hz to 200 Hz.
8. Maximal illumination of work zone, no more - 75000 lux.
9. Additional features of HTS prototype:
   a). Minimum of special requirements to operator work post accommodation;
   b). Minimal weight and size of the helmet module;
   c). Minimal price of hardware and software realized with a personal computer.

Summary (for Chapter 2)

The main methods used for development of the intelligent MMI system based on HTS and HTS+ are the following:

- coordination of natural human head movements and movements of virtual and real images of 3D objects and scenes;
- identification of spatial position and orientation of head and/or hand images in the real time mode;
- use of 3D frame models as an informative base for description of scene objects and for storing shapes of human movements;
- teaching MMI system and robot control system by showing the characteristic motions of man-operator head and hands.

Some applications of new MMI system based on the HTS prototype:
- Advanced computer interface for gesture exchange with PC;
- Pseudo-holographic effect in perception of 3D images in 3D PC displays or projective ones;
- MMI for the home (office) serving robot-like devices;
- MMI for telemedicine systems, medical robots, for rehabilitation and disabled workers;
- simulator of real time control process (nuclear station, aviation, and others);
- remote observation of environment with camera-head (Web-cameras, security etc.);
- MMI for the telerobotic control with effect of presence in remote WZ (space, UAV, UWV etc.);
- MMI for multi-robotic systems control (with two hands and head, with group of operators ets.).

Results of experimental testing MMI system with the HTS prototype, for the telecontrol with natural motions of hand as well, are presented in [13] and in Chapter 3 below.

Chapter 3. Experimental study of man-machine interface implementing systems tracking of man-operator's motions

The preliminary test results on a Man-Machine Interface (MMI) with the Head Tracking System (HTS) prototype integrated in the Hard & Software Complex (HSC) facility to achieve their appropriate performance while controlling a remote robot-manipulator are given in this chapter. The following methods and algorithms developed for the remote robot control and work zone (WZ) observation are being experimentally studied:

1. A new method for scanning and reconstructing 3D images of WZ with natural head movements;
2. Algorithms and SW for teaching robots in trajectories for automated WZ observation by showing natural motions of head or hand;
3. Algorithms and SW for automated creation of computer models and reconstruction of realistic 3D WZ images.

Besides that, a new 6D position-speed control handle design was proposed and tested for ascertaining its capability for control and teaching robot by showing natural motions of hand, based on a HTS variety (HTS+).

3.1. The HTS and HTS+ prototypes integrated in HSC facility

The efficiency of some methods was being studied in preliminary experiments: representation of 3D scenes with an effect of 3D image observation and remote control of robot by natural motions of man-operator's head & hand.

The HTS prototype integrated in HSC facility enables the following capabilities:

A). Remote Viewing (RV) of work scene using a robot-like device moving an observation cameras over a Work Zone simulating a space station fragment with payload containers;

B). Remote Control (RC) of robot-manipulator carrying the mock-up containers over the Space Station's (SS) mock-up.

The structural diagram of HSC with integrated HTS and ETS accommodated at the control post (CP) and two "PUMA" robots in the room with the space station's mock-up (work zone) is shown in Fig. 3.1 [2].

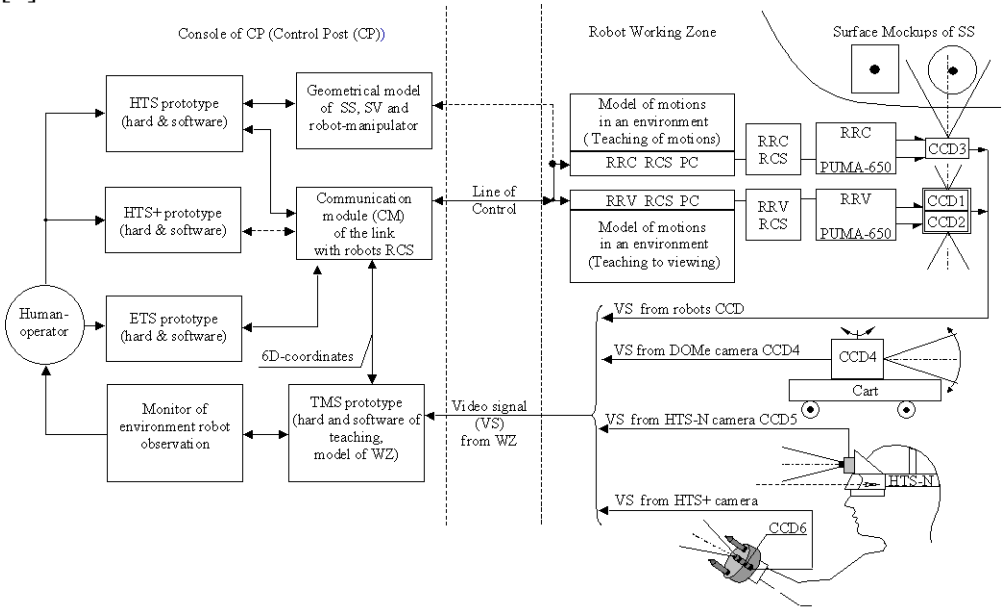


Fig. 3.1 HTS and HTS+ integrated in HSC facility for telecontrol and environment observation

This preliminary experimental study showed the efficacy of control using HTS+, especially of control with 6D position-speed handle, and, also, the possibility of co-ordinating head and hand motions.

For experimental study of distant WZ observation and telecontrol there used in HSC:
- robot-like device for WZ distant observation (Robot for Remote View - RRV);
- telecontrolled robot executing work operations (Robot for Remote Control - RRC);
- physical mock-up of a SS's fragment and containers (scaled 1:4 and a smaller mock-up 1:10) simulating WZ.

For accomplishing remote viewing and control RRV is equipped with two stereo cameras (CCD1, CCD2), mounted on the gripper for WZ observation, while robot RRC is equipped with one camera (CCD3) for visual controlling manipulation with objects (containers).

Coordinate control data from HTS and HTS+ are transmitted via communication module CM (serial interface) to two control systems (RCS), those of RRV and RRC.

HW and SW of Television Measurement System (TMS) [9] are used for receiving, processing, viewing and 6D registering the video image with computer models. The processed actual WZ image registered with WZ computer model (GM) is graphically displayed on the control post PC monitor.

Besides that, the following additional equipment is used for experimental study of different methods for distant WZ observation:
- panoramic camera a DOOM type (CCD4), may be , also, mounted on a movable cart;
- 2 cameras, on head (CCD5) and hand (CCD6) of man for simulating astronaut's work beyond a space station (SS), and, also, for teaching RRV in movements for WZ observation.

During this stage of the experimental study the robots-manipulators were controlled separately: RRV with HTS prototype and RRC with the HTS+ prototype.

Experimental study of robot telecontrol using HTS and HTS+ was performed preliminarily with geometric models (GM) of the space station and the robot-manipulator. As was noted in [2], this stage is necessary prior to operator's controlling a real robot for testing operability of HTS prototype.

This test (prediction of results) is to be done for ensuring safety of expensive equipment and, also, for finding an optimal trajectory of robot-manipulator's work tool amidst environmental objects.

Execution of a trajectory may be realized by robot-manipulator in an autonomous mode guided by a model of movement stored in RCS of the remote robot. This model is not necessarily transmitted to the robot via communication link in real time this may be done prior to executing it.

3.2. Remote environment observation with the robot-like device controlled by HTS

3.2.1. General scheme of the remote RRV viewing with HTS prototype

The remote WZ observation mode was studied with the space station mock-up and payload containers (Fig. 3.2). Observation cameras CCD1 and CCD2, mounted on the robot-like device's (RRV) gripper, transmit video signals of images of WZ objects. The RRV, is controlled with HTS from the control post (CP) via the communication link with the Robot Control System (RCS).

The WZ video is transmitted from cameras (CCD1-CCD2) to the CP and displayed in stereo mode on PC monitor. The coordinate information on cameras position in WZ space goes from RCS through communication link to the CP PC at Control Post.

Man-operator at CP with natural head motion controls the RRV borne cameras at remote WZ.

The principle of remote WZ observation using HTS consists in the following: the operator is sitting at CP before the PC-monitor with the reference device (RDU) being built in the head garniture.

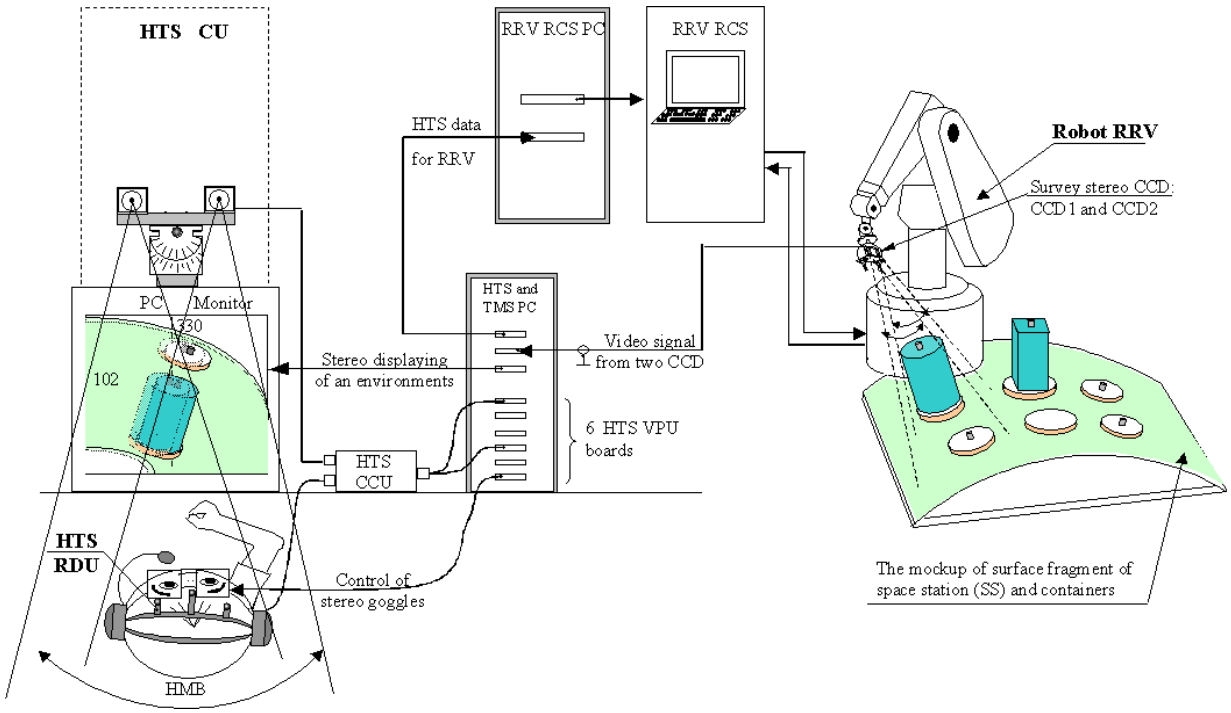


Fig. 3.2 General scheme of the remote RRV viewing with HTS prototype

The operator's head turn is perceived with cameras (CU) of HTS mounted immovably on the monitor and there images processed in a video processor (VPU) at PC (HTS&TMS). The resulting head position data are passed to RRV RCS where they are transformed into the robot coordinate system and converted into RRV control signals. As the result, RRV performs the head-commanded observation movement.

The images of cameras are CCD1, CCD2 transmitted to the control post computer which displays WZ stereo picture on the monitor the operator sits at. Equipped with the stereo goggles, the operator observes 3D WZ images and may examine them in more detail because turning and approaching head to PC monitor changes point-of-view and scale of observed objects.

3.2.2. Specific features of scanning a real scene using a camera borne by a robot-like device

The following real scene scanning modes were studied:

- observation with RRV-borne cameras controlled with HTS;
- control of camera movements with hand or head;
- autonomous tracking objects with a camera following trajectories after teaching by showing;
- telecontrol with elements of prediction with GM;
- telecontrol using the Augmented Reality principle, i.e. supplying the real video with GM or otherwise;
- operation with an augmented computer model supplied with real video (current or earlier recorded).

The experiments had been carried out with a camera mounted on head (helmet) and on hand (bracelet) of man-operator (Fig. 3.3). These modes are important for simulating man activity at remote WZ (e.g. astronaut in outer Space).

a).Camera mounted on helmet b).The image from camera on helmet c).Camera mounted on bracelet d).The image from camera on bracelet

Fig. 3.3 Scanning with movements of head and hand

The difference in head and hand scanning modes is determined by difference in arm and head kinematics and more convenient access by hand to hard-reached WZ places. The performed experiments have established the expedience of using these two different modes of scanning. By the head-borne camera scanning with speed to higher than 40 ang.deg./s the moving image is perceived as naturally coordinated unlike that obtained with the hand scanning. In the latter case the naturalness of perception is attained at far lower angular speeds (below 10 angle deg./s) and requires a strictly defined trajectory of hand (analogous to that of head), e.g. controlling camera's rotation round its centre or uniform movement over a circle or line.

Using operator's natural movements of head for scanning WZ is caused by a necessity of making his hands free for work operations.

3.2.3. HTS application for remote WZ observation using a robot-manipulator

The difficulty of image perception for man-operator (in CP) consists in a fact that observational movements of other operator (in WZ) do not conform with his own movements both in time and space, especially while operating in real time mode.

While scanning WZ with a robot the situation is simpler because the robot does not move the camera autonomously but slaves the head of operator (with control from HTS). On the other hand, the WZ operator's situation is more easy for he may take a right way at his will, especially in emergency situations, and transmit the picture to PC operator. Therefore, it is necessary to create a MMI enabling coordination of CP operator's perceiving WZ both with a man-operator or robot.

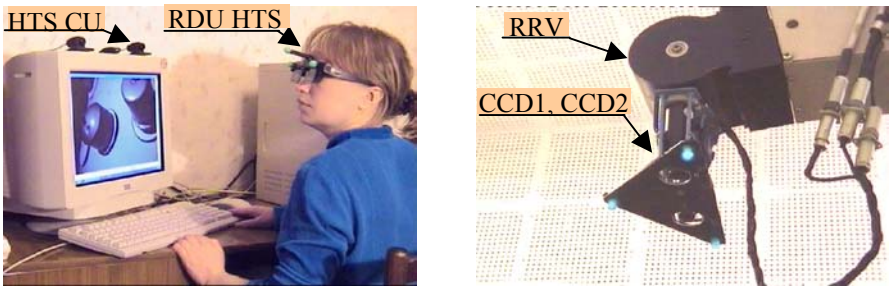
In both methods scanning with a narrow-field CCD camera takes lesser fragments of WZ but with lesser distorsion than a wide-field camera. In the former case a longer observation procedure is added to a long process of making the entire picture of WZ (panorama synthesis), in the latter case we obtain a full picture but with a strong distorsion (needed to be corrected for). Running a CCD camera over trajectories with known coordinates makes splicing of fragments into a unified panorama more easy. Otherwise, one is compelled to identify common elements for references in spliced fragments for their integration in a panorama [13].

HTS (HTS+) aided telecontrol of camera movements will enable help of CP operator's mind and his intuition which reveals itself in natural reactions while choosing inspecting object of interest Fig. 3.4 [14].

Fig. 3.4 Using HTS for robot-manipulator control

Thus, using the systems tracking movements of man enables video data representation in a form the most convenient for CP operator and excludes the nonconformity of WZ camera movements to those desirable for PC operator, what is confirmed by the experimental data. For acquisition of CCD1 and CCD2 positions relative the SS mock-up a reference device unit (RDU) is used, and the HTS camera is immovably mounted on the SS mock-up like HTS's CU.

An operator using HTS at the CP controls the position and orientation of RRV gripper with cameras mounted on it. The resulting picture he sees on the PC monitor (see Fig.3.5).



a). Operator at the control post b). General view of the WZ mock-up with the robot-like device RRV

Fig. 3.5 Prototype of equipment for remote WZ observation using HTS

3.2.4. Structural scheme of the algorithm for obtaining, processing and displaying Work Zone

WZ remote observation may be accomplished also with one camera mounted, e.g., on RRC [2]. The control algorithm with a robot for viewing different fragments of scene (or different objects) provides high scanning accuracy and is realized for earlier learned inspection trajectories or by operator's immediate hand control (HTS+) (see Fig. 3.6).

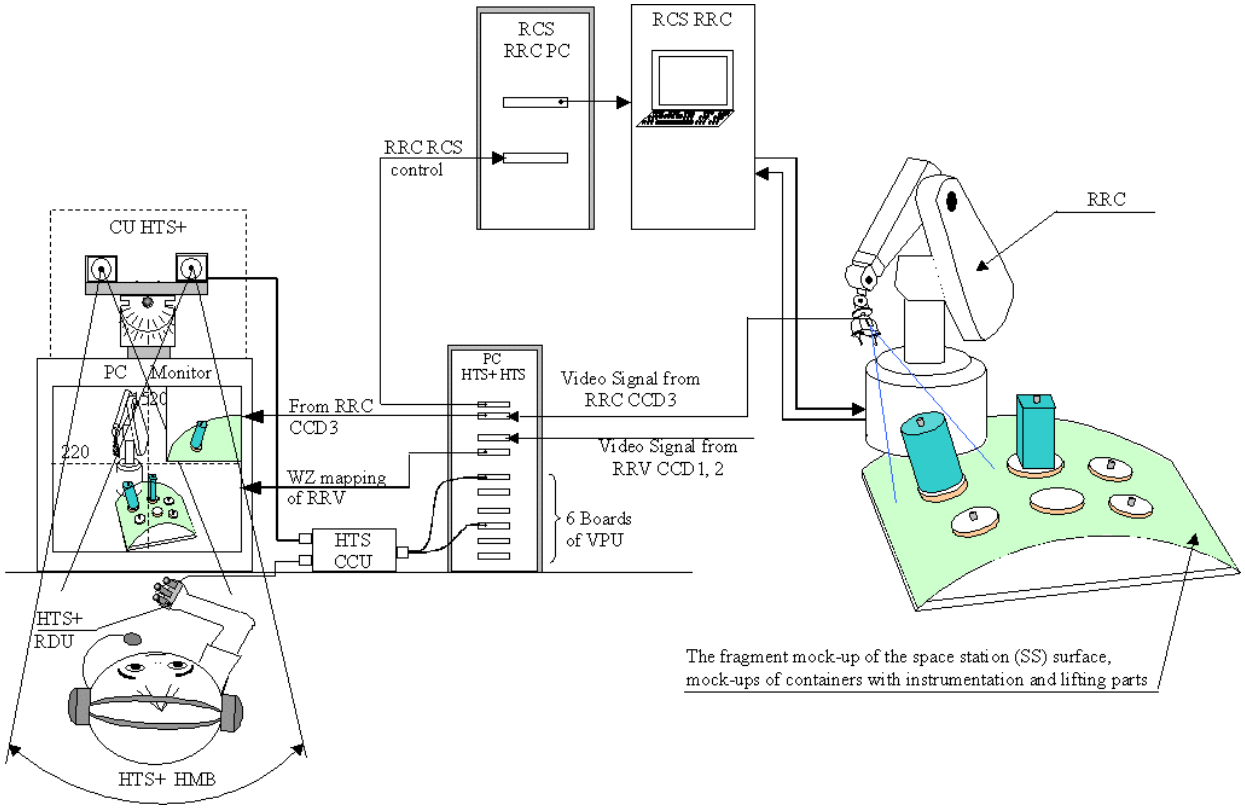
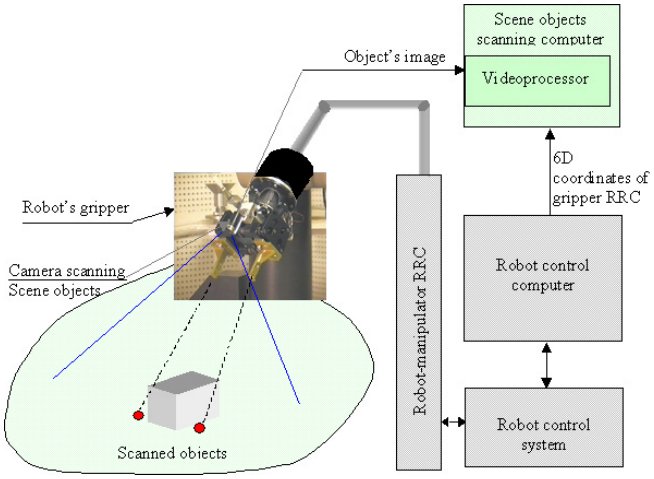
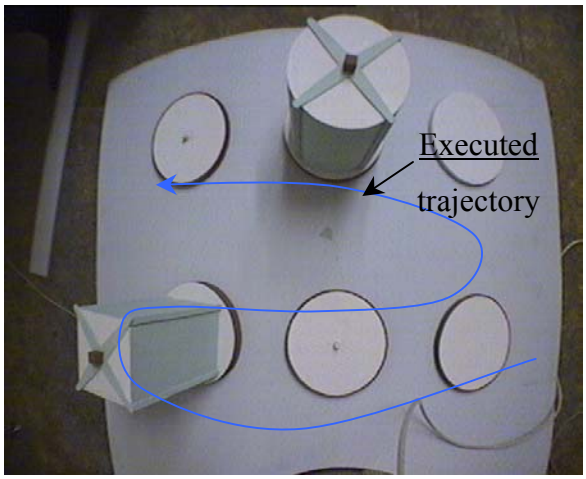


Fig. 3.6 Functional diagram of telecontrol using HTS+

The trajectories for inspection of objects or work scene take in account not only scene objects' configurations but, also, conditions for better automatic identification of image specific features and making use of results of previous measurements. Scanning WZ with a HTS+ controlled robot-manipulator RRC is accomplished by moving over an earlier determined and learned trajectory. The structural diagram is given for a prototype of telecontrolled WZ scanning system using a single CCD3 mounted on the gripper (Fig. 3.7).



a). Structural diagram for scanning scene with a robot-manipulator



b). The scanning scene of SS mock-up

Fig. 3.7 Scanning scene objects employing a robot-manipulator and HTS+

While WZ scanning with a robot-borne camera the operator views the monitor picture without any stereo goggles, it is important in a long work. A 3D picture may be reconstructed yet camera being single by usage of special algorithms.

It is realized for two consecutive images taken from two points of a known trajectory segment. The algorithm for 3D scene acquisition for a single camera is given in [4].

Single-camera robot scanning with HTS+ control was tested experimentally and proved its operability while reducing twice hardware and less straining the communication channel and computer resources in image processing. In dealing with visual data obtained in the work scene of robot-manipulator three basic modes of operation may be distinguished in the algorithm for visual representation of remote WZ (Fig. 3.8).

Mode I - remote telecontrolled surveying (scanning) of robot-manipulator work scene for acquisition of 3D video data. The survey is accomplished before the process of real robot telecontrol and serves for giving operator knowledge of work scene and manipulation objects. Meanwhile the operator sits in the remote control post. Therefore, he needs visual data to have adequate knowledge of work scene that may be obtained with cameras differently placed in the work scene. E.g. 2 cameras mounted on RRV robot give knowledge of work scene and one mounted on the gripper of RRC robot images manipulation objects (see above Fig. 3.5b and 3.7a).

Mode II - observation in the process of robot telecontrol. The cameras on RRV and RRC are controlled using HTS and HTS+ system by motion of head and hand for watching work of RRC robot. Now, Mode I is used for real time displaying work scene visual data and stabilization of real images to help operator in observation of work scene. Ability, also, is provided for registering real and GM images of work scene.

Mode III - automatic tracking at the camera telecontrol. Automatic continuous tracking RRC robot movements and manipulated objects in its gripper by cameras in RRV gripper makes possible to eliminate hand control tracking routine. All the three algorithm modes may in some cases work together, e.g. for accomplishing a hard task of simultaneous WZ observation and telecontrol of two robots RRC&RRV and, also, in perspective for automatic GM generation.

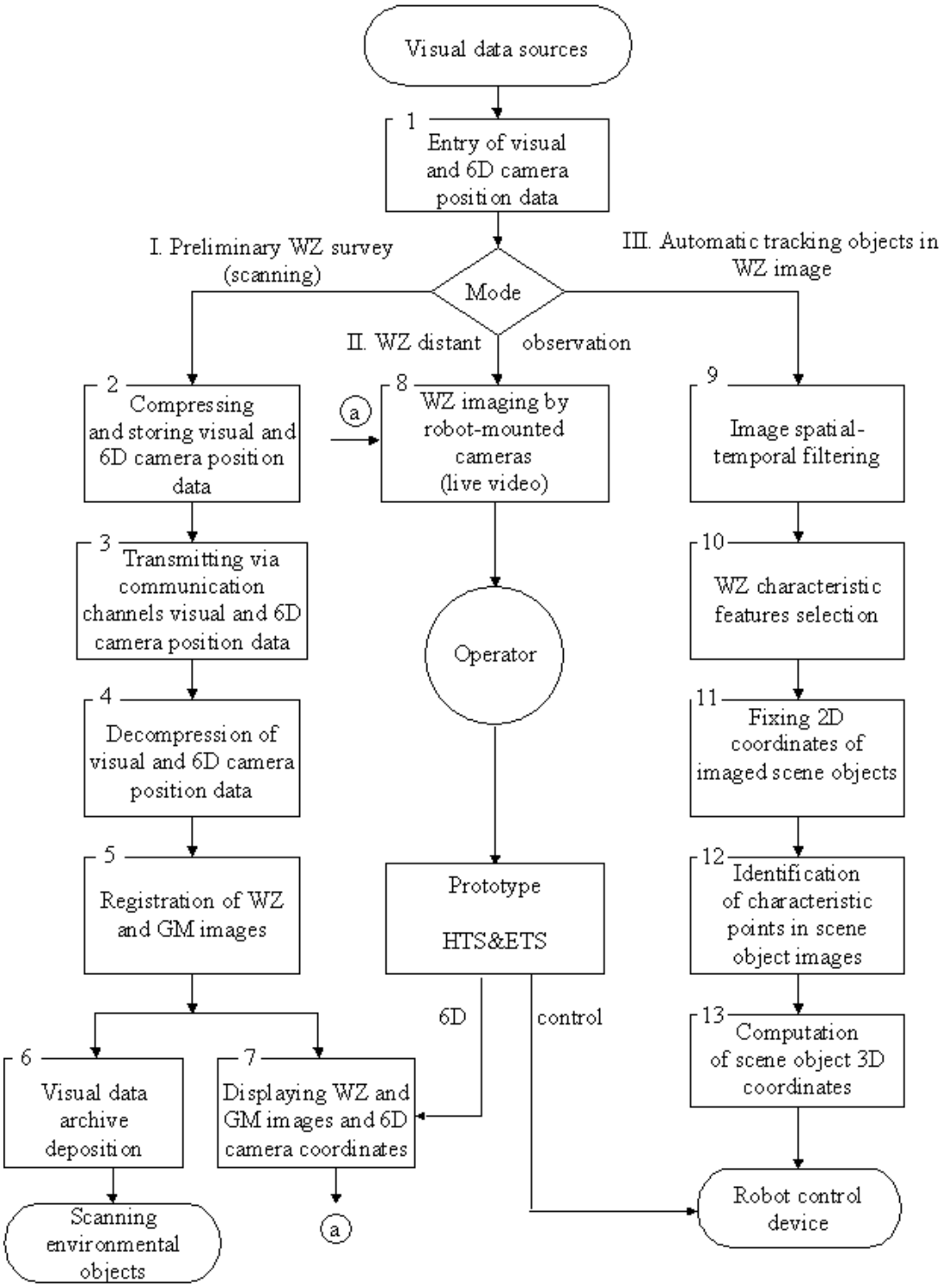


Fig. 3.8 General algorithm of WZ observation by robot-like device

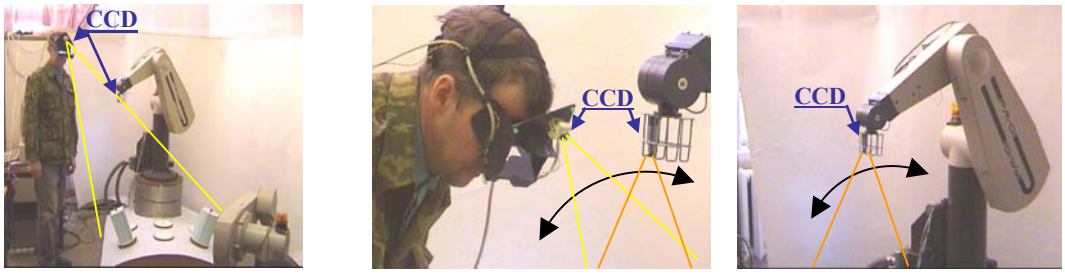
3.3. Development of hardware and software prototypes for teaching robot - manipulator by showing natural motions of operator's hand

Experimental studies are carried for teaching a robot-manipulator in trajectories by way of showing. They include experiments in teaching by showing trajectories of scene observation (using HTS) and those for getting the robot's gripper in a position for taking object with the help of HTS+.

Besides that, basing on a methodology studied earlier with the simplest objects [6, 10] and on algorithms with GM [15] the experiments were performed in the following procedure:

1. Man-operator seeing a stereo picture obtained by RRV robot makes observational movements of head taking in account kinematical characteristics of RRV;
2. Panoramic picture tagged with coordinate data is stored in Graphic Station at CP, upon smoothing of a trajectory with a moving camera model and earlier stored panorama of environment;
3. RRV control with signals following the memorized trajectory is enabled, head motion disabled, man-operator verifies the fidelity of reproducing the taught-in movement.

The experimental study of the method for robot-manipulator's control using HTS (Fig. 3.9) and HTS+ (Fig. 3.10) was conducted with the models (physical and virtual) of objects and also teaching the robot by showing movements [7].



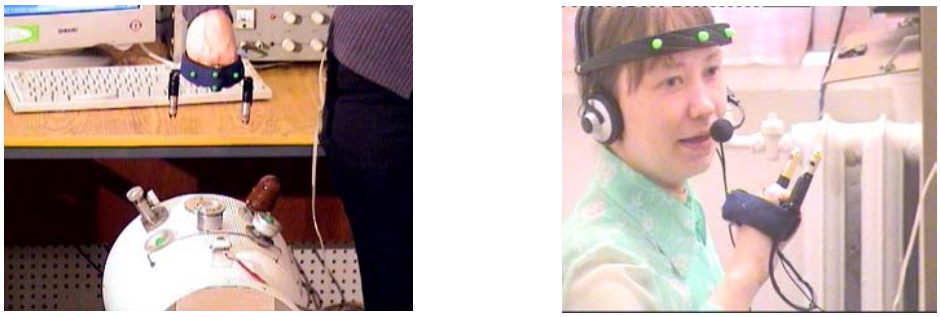
a). Teaching RRV in WZ observing movements b). Teaching by showing process c). Robot executing learned movements.

Fig. 3.9 Robot teaching by showing with HTS

The SS mock-ups scaled 1:4 (Fig. 3.9 a) and scaled 1:10 (Fig. 3.10 a) were fabricated for teaching robot by showing. The mock-ups are easily accommodated in operator's work place and enable effective teaching both robot and the man-operator himself.

The man-operator observes WZ with HTS & HTS+ using his own experience in it. Robot RRV repeats the procedure complying with the order following the trajectory of man's head. And prior to that RCS of RRV may move the cursor over the obtained panorama image, vary the scale of fragments and, upon ascertaining all being right, execute the actual movement.

Adding an HTS+ to the above system will provide a natural coordination of hand and head control movements: head controls the observation camera, mounted on a special robot-like device RRV and hand controls position and manipulations of the main robot-manipulator (RRC) (Fig. 3.10).



a). Teaching by showing with HTS+ b). Control by HTS&HTS+

Fig. 3.10 Telecontrol and teaching of robot using combination of systems "head & hand"

While using position-speed HTS + 6D handle for control of robot - manipulator one gets a larger range and naturalness of movement than those attained with a traditional handle of "Master-Arm" type (Fig. 3.11) [16]. The design and appearance of the HW for telecontol are the same as for remote observation. Difference is in using the position-speed control handle with RDU and another robot (RRC) manipulating the container mock-ups.

a). Position-speed HTS+ 6D handle for robot control b). Controlled robot (RRC)

Fig. 3.11 Position-speed HTS+ 6D handle for robot control.

3.3.1. Designs of the position - speed control handle for operation with HTS+

The experimental study of the MMI for robot-manipulator's telecontrol using HTS+ was conducted with the RDU bracelet using models (physical and virtual) of objects and, also, teaching robot by showing movements. Within the framework of this Project some variants of RDU design for HTS + were developed (Fig. 3.12):

1). A bracelet with passive reference marks, laser pointers and miniature CCD camera that the space station's mock-up is observed with (Fig. 3.12 a),
2). Bracelet with active reference marks (IR-LEDs) (Fig. 3.12b),
3). Hand-manipulated trident-like device with IR LEDs (Fig. 3.12c),
4). Articulated arm, with the bracelet on it, capable of fixing its position (Fig. 3.13).

a). b). c).

Fig. 3.12 Variants of RDU design for HTS+ prototype

Development of a new type of control handle was stimulated by a fact that operator holding RDU in his hand will get tired to keep it long time without support of hand. Therefore the RDU on handle was fixed on a special supporting mechanism (articulated arm) capable of RDU fixation in desired position (see Fig. 3.14 b below). In the course of experimental studies of telecontrol with HTS+ a new design of 6D handle for position and speed control was proposed, one based on employment of HTS' HW and SW. A structural diagram of the position-speed control handle on the base of the HTS prototype is shown in Fig. 3.13.

A bracelet with RDU, whose position and orientation are tracked with HTS+, is mounted on the last link of a 6 DOF multi-link mechanism (articulated arm). This mechanism has 5-8 joints with self-fixation by friction in a motion mode and handle for hard fixation in a stand mode. For providing simultaneous control for speed the last joint of the arm is equipped with a force & torque sensor [15].

The last link of this mechanism is stiffly coupled with a force-torque sensor enabling operator to 6D speed control of robot-manipulator gripper movements. Thus the gripper is controlled for position and orientation with data from HTS+ (x, y, z, φ_x, φ_y, φ_z) and for speed with data from the force-torque sensor ($\dot{x}, \dot{y}, \dot{z}, \dot{\varphi}_X, \dot{\varphi}_Y, \dot{\varphi}_Z$).

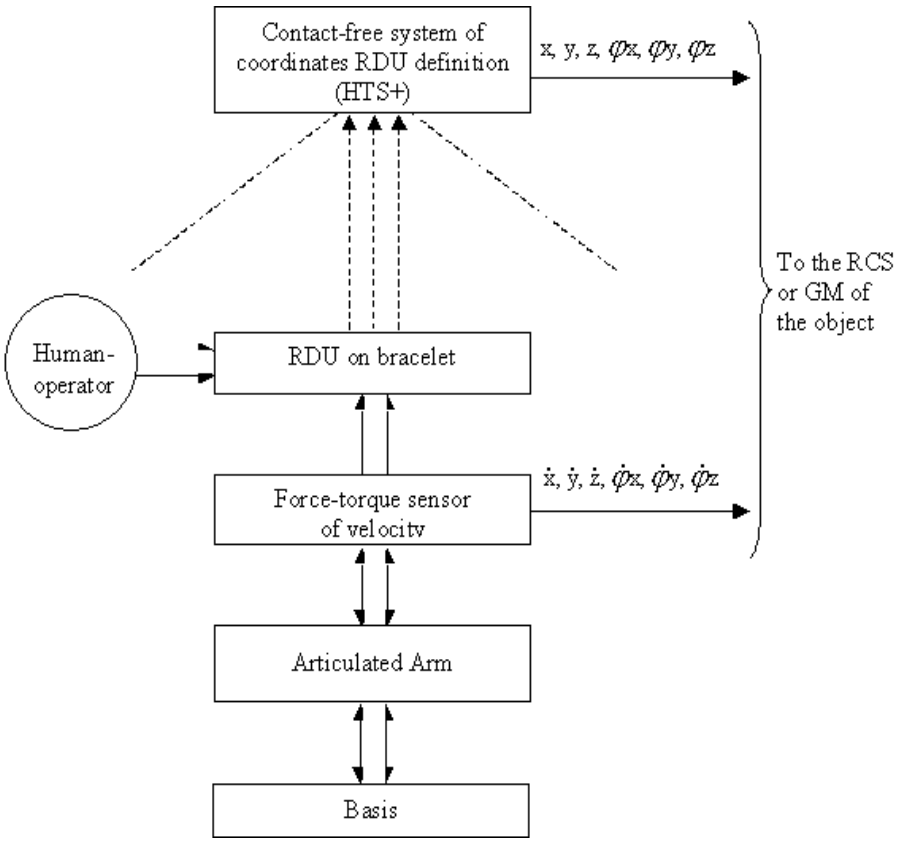


Fig. 3.13 Structural diagram of HTS+ position-speed 6D control handle

The arm's base is fixed on operator's table in front of the monitor with mounted on it cameras. The CU of HTS+ detects the RDU position and orientation converting them in 6D control signals for a computer model or robot's RCS. The arm's operating range depends on length of the fixation mechanism and may be extended to all volume of operator's work place. The HTS+ handle design for the new version differs from its prototype by a more convenient operator's handling (all advantages of the pilot handle are preserved) and more robust design. A design of the HTS+ 6D handle prototype with articulated arm and RDU is shown in Fig. 3.14b.

a). The 6D handle with force & torque sensor

b). New HTS+ 6D handle ("Cober")

Fig. 3.14 The new 6D position-speed control handle HTS+

A salient feature of the proposed design is a combination in one control handle of speed and position assignment [Russian patent]. A speed control mode is realized in a fixed arm position using a 6D force-torque sensor. The developed new design of the handle (Fig. 14 b) for its basic parameters is similar to known analogues [17, 18], yet it surpasses them in mass, dimensional, accuracy and a low price (see Table 3.1).

The main advantages of the proposed HTS+ 6D handle design are the following:
- relatively large zone of free motion of hand (radius of the zone from 300 to 600 mm);
- the least attainable inaccuracy of positioning (about 0,1 mm and 0,2 angle deg.);
- minimal mass and dimensions of the system as a whole (no more 1 kg and 300×200×100 mm^3);
- low price of serially produced article (about $ 800).

The new 6D handle may be fitted for individual user hand, ensuring with it stability of the initial position while controlling a robot.

Table 3.1

Comparison of HTS+ handle with other types for robot control

Parameter name	Handle type					Note on comparative advantage
	Two «Souz» handles	Pilot Force-torque	«Delta» force torque	robot-like handle	HTS+ handle	
Degrees of Freedom (DOF)	3+3	6	6	6	6	=
Hand movement coordination	Difficult	Simple and natural	Simple and natural	Simple and natural	Simple and natural	Simple and natural
Speed control	has	has	has	has	has	=
Position control	In limited zone	has	has	has	has	=
Hand Motion Box (HMB+), mm	50	-	150	1500	300-600	Unrestricted volume comfortable for operator
Speed of movement, deg/s	>200	-	>30	>30	>200	High
Force-torque feedback	-	-	has	has	possible	Possible in perspective
Positioning accuracy, mm	-	-	1	0,1	0,1	=
Position fixation	-	-	has	has	has	=
Data sampling rate, Hz	≥ 1000	≥ 100	≥ 1000	< 20	> 100	Medium
Supply voltage, V	27 DC	27 DC	220 AC	220 AC	12 DC	Minimal
Power consumption (without PC), W	5	10	200	> 1000	5	Minimal
HW weight (without PC)	10 кg	3 кg	15 кg	> 200 кg	1 кg	Much lower
Interface with PC	RS-232	RS-232	RS-232	RS-232	RS-232	=
Design kind	Portable	Portable	Mobile	Stationary	Portable	Portable
HW dimensions (without PC), mm	200×100 ×400	200×200 ×200	800×600 ×600	2000×2000 ×2000	300×200 ×100	Minimal
Serial production	Possible	Possible	Possible	Very difficult	Possible	Possible
Novelty	Absolute	Up-to-date	Novel	Up-to-date	Novel (patented)	New
Control force	Minimal	Minimal	Medium	Large	Minimal	Small
Cost (thousand USD)	3,5	1,2	12,0	> 50,0	0,8	Minimal

A disadvantage of HTS+ handle, as compared with the HTS+ bracelet, is absence of freedom for fingers to execute operations or be simply at ease. Besides, mass and dimensions of the new handle are more large as compared with the bracelet, especially, when the marks on it are passive ones.

The experiments have shown that for precise control of position and orientation (with inaccuracy less than 1 mm and 1 ang. deg.) the new HTS+ handle is the most effective.

The HTS+ bracelet is very effective for robot teaching by showing direction of movement and gives natural freedom to hand while robot tracks it.

An additional advantage of this design of the handle is a possibility of mounting on it knobs and the tracking boll for a finger control of fine robot gripper movements.

Summary (for Chapter 3)

The operation of HTS and HTS+ with the space station mock-up and actual robot-manipulators is realized. Experimental studies of the algorithms and SW for GM control had been carried [19]:

A. The position and speed control with natural movements of head and hand was studied.
B. Limits were established of the zone of stable control for linear and angular movements of RDU. Large ranges of hand movements were demonstrated and necessity was established of hand position fixation for ease of control.
C. Experiments were accomplished for estimating control process dynamics in various operation modes. Results were obtained that show a reduction in dynamics when HTS and GM SW are operated on one computer and the necessity of two computers (host PC and graphic station) for the control algorithm's realization was established.
D. The efficacy of passive HTS and HTS+ was tested while working with color reference marks on the actual background of the control post (CP), with sunlight reflected from walls and with lamps appearing in the camera's FOV.

Chapter 4. Novel man-machine interface for telerobotics using eye tracking systems

Preliminary results on a Man-Machine Interface (MMI) development for the robot telecontrol basing on man-operator gaze tracking system are proposed. The 3D scene representation with the 3D virtual cursor control and the remote robot control by means of natural motions of man-operator gaze has been studied.

The main part of this chapter is devoted to the components of the Eye Tracking Systems (ETS) and their integration into MMI system. The optical scheme of the ETS prototype for gaze direction acquisition is described too.

One of the most promising applications of gaze pointing for robotics is developing the idea of the 3D "virtual cursor" (VC) pointing used in systems of Augmented Reality (AR) and stereo vision. It provides high dynamic and natural easiness in virtual cursor pointing.

The proposed methods for ETS control of High Resolution Image Zone (HRIZ) enabled a considerable reduction of TV pass-band without spoiling the perceived sharpness of picture.

4.1. ETS hardware prototype composition

This work proposes a novel MMI having the following features:
- High dynamic interface with computer systems by means of the natural human gaze tracking;
- Realization of a high fidelity image with gaze controlling a local high-resolution zone in displayed image for usage narrow pass band communication channel;
- Simplicity and naturalness of the robot teaching process by means of showing natural human gaze motions.

The control by gaze is realized with combined Eye Tracking System (ETS) and Head Tracking System (HTS) for control of spatial position and orientation of images or real dynamic objects.

Basing on optical-television methods for eye tracking the ETS prototype is developed in addition to traditional computer manipulators of a mouse type. The ETS prototype consists of the following units (similar to a structure of the HTS prototype), see the structural scheme (Fig. 4.1) [1, 2]:

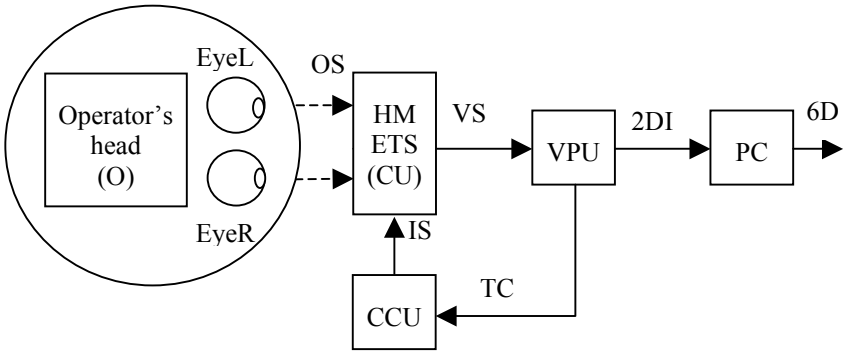


Fig. 4.1 Structural scheme of the ETS hardware prototype

- Camera unit (CU) with two b/w cameras mounted on the helmet module (HM ETS);
- Camera control unit (CCU) for power supply, pulse control of eye illumination and cameras exposition;
- Video processor unit (VPU) for digital processing of the camera video signals;
- Personal computer (PC).

IR LEDs in HM illuminate the left eye and right eye separately. Two CCD cameras in HM ETS detect optical signal (OS) reflected by eyes and produce video signal (VS).

Video signals VS from cameras are inputted to the VPU, which realized a differential processing of eye images. In the process of the ETS prototype perfection, the major part of processing has been realized with VPU hardware means. The data of eye image (2DI) processed by VPU are transferred to PC.

To realize basic functions of image processing and computation of eye rotation and convergence angles (6D) the prototypes of active and passive ETS utilize hard & software means realized, accordingly, at the video processor unit and PC Pentium 3 (4).

4.1.1. Main principals of ETS prototype operation

A new method for realization of the prototype is based on measurement of increments in angular coordinates of the eyeball axis relative to some known position of this axis obtained by independent way. The parameters to be measured in this method are angular coordinates of only one element – corneal reflex. This method requires a more complicated algorithm and calibrating procedure for acquisition of angular coordinates of the eyeball axis in one or several reference points with reasonable periodicity in the process of operation (Fig. 4.2).

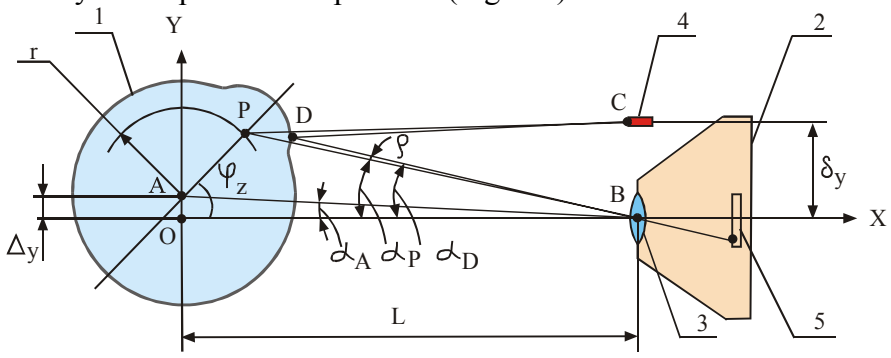


Fig. 4.2 Optical scheme of ETS for determination of look direction

In the scheme there are shown eye (1) with the center of rotation in point 0, video camera (2) with the lens (3) and a point source IR LED (4). Light rays issuing from IR LED point source (4) illuminate eye 1, reflect from point D on the cornea, and come to the camera's FPA (5).

Angular direction to the corneal reflex fixed by the camera is determined as angle α_D between segment BD and the camera axis OB. Angular direction to the corneal center of curvature P is determined for camera axis AB and differs from α_D at value ρ little depending from value of angle α_P and determined with an approximate expression: $sin\ \rho = R\delta_y/2\ L\ (L - R)$.

For an average eye one takes R = 8 mm. Allowing, for example, L = 80 mm and δ = 8 mm we will have: $sin\ \rho$ = 0,0056, $\rho \approx 0{,}32° \approx 20'$. So, for measured value α_D, α_P may be determined from expression:

$$\alpha_P = \alpha_D - \rho \qquad (4.1)$$

Let us designate angle of rotation of the eye axis AP in the vertical plane respecting the camera's axis OB as φ_z (round axis OZ) and angular direction at the eye's center of rotation A as α_A (angle between segment AB and the camera axis OB). Solving the triangle APB we will obtain expression:

$$sin(\varphi_Z + \alpha_P) = L\ sin(\alpha_P - \alpha_A) / r\ cos\ \alpha_A \qquad (4.2)$$

Taking in account (4.1) we get:

$$sin(\varphi_X + \alpha_D - \rho) = L\ sin(\alpha_D - \rho - \alpha_A) / r\ cos\ \alpha_A \qquad (4.3)$$

The abovementioned considerations for vertical angles φ_Z are just and for horizontal angles φ_Y (round axis OY) of eye turn with this difference that displacement of LED's 4 emission center C from axis OB of the camera along axis OZ may be none, i.e. δ_Z=0. Then, ρ=0 and $\alpha_P = \alpha_D$ and we get:

$$sin(\varphi_Y + \alpha_D) = N\, sin(\alpha_D - \alpha_A) / r\, cos\ \alpha_A \qquad (4.4)$$

Parameters L, ρ and α_A are determined by the helmet module's position relative to eyes and should be established in the process of initial calibration by solving a system of equations (4.3) - (4.4) for enough big number of pairs (φ_Y, φ_Z) measured with independent from ETS method, e.g. with HTS. Upon computation of parameters L, ρ and α_A and writing them to the read/write memory of ETS processor the current values (φ_Y, φ_Z) will be continuously computed by ETS with formulas (4.3) and (4.4).

Thus, the method enables determination of operator's gaze (φ_Y, φ_Z) in the camera system of coordinate and, hence, in the helmet-bound coordinates system $O_HX_HY_HZ_H$. For attaining a required accuracy the cameras of ETS are calibrated.

4.1.2. The design of helmet module (HM) of the ETS prototype

For studying the gaze-measured method experimentally an ETS helmet module hardware prototype was developed and fabricated, the design scheme of one version is shown in Fig. 4.3.

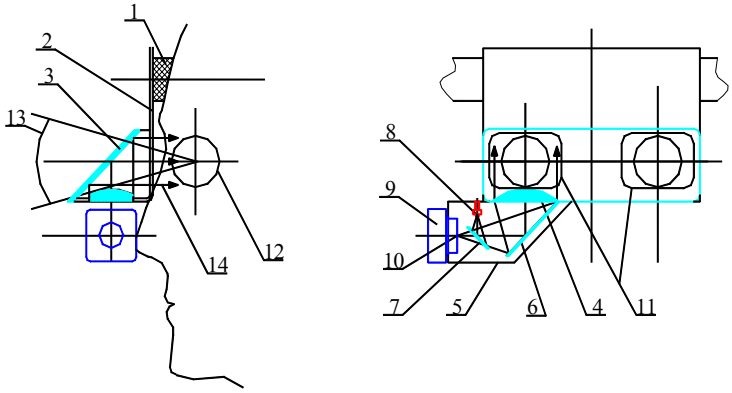


Fig. 4.3 The variant of the ETS prototype's HM design scheme

A bracket 2 is fixed to the headpiece 1 bearing a beamsplitter 3 and a collective lens 4. A housing 5 is mounted on the bracket in which a mirror 6 and a semitransparent mirror 7, an IR LED 8 and a CCD camera are accommodated.

The entrance pupil center 10 of the camera lens and the center of LED emitting area are optically conjugated with the help of semitransparent mirror 7 and combined with the focus of collective lens 4.

The bracket 2 has rectangular holes 11 in it through which eyes 12 of operator view via beamsplitter 3 the environment. For raising effectiveness of operation in conditions of ambient illumination and providing undistracted viewing of environment the following requirements to ETS should be satisfied:

- beamsplitter 3 must have a selective coating maximally reflecting IR LED light with transmitting in the visible (400…700 nm);
- semitransparent mirror 7 must be coated for 50% transmission and 50 % reflection of IR LED light by minimal absorption;
- CCD camera lens should be provided with a built-in spectral filter having minimal transmission in the visible (400…700 nm) and maximal for IR LED emission;

Angular field of view 13 should be no less than 30º in vertical plane and 60º in horizontal plane.

The ETS prototype operates in the following way. Divergent rays emitted by IR LED 8 successively reflect from mirrors 7 and 6, come to collective lens 4 and go out as a parallel beam 14 which being reflected from beamsplitter 3 illuminates the area of possible position of the eye elements, reflects from beamsplitter 3, comes to collective lens 4, reflects from mirror 7 and is focused at the entrance pupil of camera lens 10.

Light caught by the aperture forms on the camera's FPA optical images of the eye pupil and cornea reflex. A stray light from off-axial sources is also possible. Video signal from the camera comes to the PC where it is processed by a specially developed program and values are calculated of angular position of the eye optical axis relative to the system of coordinates bound with the helmet-mounted unit.

The design of the ETS prototype's Helmet Module (HM) and Head Mounted Display (HMD) is shown in Fig. 4.4.

a). HM ETS with two optical channels b). HM HTS with Cy-visor DH-4400 HMD

Fig. 4.4 Design of HM ETS and HM HTS with HMD

4.1.3. Technique of ETS calibration using HTS

Two calibration modes for the ETS prototype are supported:
- initial calibration including fitting for individual operator ranges of eye turn angles to the monitor image size (Fig. 4.5a);
- ETS calibration within image area using HTS data (Fig. 4.5b).

a) preliminary calibration of ETS b) calibration of ETS by HTS

Fig. 4.5 Calibration procedures of ETS

Let's consider one of possible versions of calibration technique for ETS operating with an optic-television HTS and monitor displaying the work scene [11].

Images of points are displayed on the monitor. Near each point on the screen a mnemonic tag helps to the operator to put HM in the required position relative to the monitor and HTS' CU for five coordinates: two linear (Y,*Z*) and three angular ones (φ_x, φ_y ,φ_z). Distance to the monitor (coordinate X) in this case has no importance.

In the same time as operator fixes look at each of points, the ETS determines the corresponding angle values for left and right eyes. The corresponding pairs of angle values are determined with HTS and the calibration parameters of ETS are calculated.

For typical distance to the monitor L=400mm and eye base B=66 mm the value of calculated angle φ_i will be:

$$\varphi_i = arctg\,(B/2L) \div arctg\,(B/L) = 9° \div 18°$$

During the calibration process the HTS operates practically in stationary mode in the central zone of operating angles. In such conditions the HTS provides measurement of HM position relative to the monitor screen with an error $\Delta\varphi_{HTS}$=5′-10′ for angular coordinates and ΔL_{HTS} = 2-3 mm for linear ones, what will result in 0,5÷1% for L and φ_I. Thus, it is acceptable as $\Delta\varphi_I \approx 8'$.

The ETS operates also in stationary mode and in the central zone of operation angles without external light interference, therefore it is capable to measure α_{Di} with maximum accuracy attainable for this method: $\Delta\alpha_{Di} \approx 10'$.

Besides the instrumental components $\Delta\varphi_I$ and $\Delta\alpha_{Di}$, it is necessary to take into consideration «the human factor»: the operator can fix look at point with a simular error: $\Delta\varphi_O \approx 12'$.

The total error Δ_Σ of this method can be defined as the root of sum of square of three errors - $\Delta\varphi_I$, $\Delta\varphi_O$ and $\Delta\alpha_{Di}$:

$$\Delta_\Sigma = [(\Delta\varphi_I)^2 + (\Delta\varphi_O)^2 + (\Delta\alpha_{Di})^2]^{0,5} = [8^2 + 10^2 + 12^2]^{0,5} \approx 17'.$$

4.2. Algorithms and software of ETS prototype

The realization of the ETS prototype involves measurement of 2 components (φ_z,φ_y) for 3 rotation angles of each eye in the head coordinate system.

The ETS prototype must satisfy the following requirements:

1). Measurement must be provided of eye turn angles in limits maximally close to the natural oculomotor range (about 120° for yaw - φ_y and 60° for pitch φ_z;
2). Eye turn measurement error for angles (φ_z, φ_y) must be no more than 0,1° and must not depend from some displacements of the helmet;
3). System operability must not depend from colour of eye iris and size of the pupils;
4). Protection must be provided against interfering illumination at the control post while creating the illumination to 75000 lux;
5). Helmet-mounted part of the system must not prevent seeing visual information displayed at the helmet monitors and, also, observing surrounding real scene;
6). Helmet-mounted part of the system should have minimal weight and dimensions.

4.2.1. The general operation algorithm for the ETS prototype

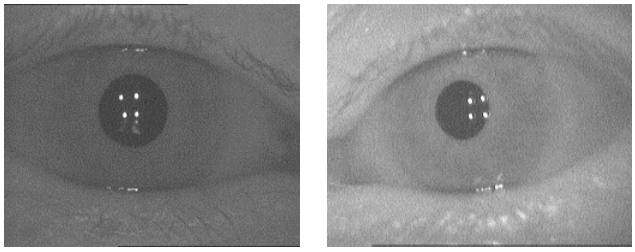
The general operation algorithm for the ETS prototype, common for active and passive varieties, consists of the following operation steps:

1). Initial adaptation to operator's eye (alignment and adjustment for individual operator);
2). Spatial-temporal filtering of eye image zone;
3). Selection of eye image specific features (pupil, iris and eye corners);
4). Identification with a model of Eye image specific features;
5). Sub-pixel measurement of coordinates of eye image elements: cornea reflex, etc.;
6). Computation of turn angles, eye convergence and torsion rotation angles;
7). Measurement of eye speed vectors (saccade), drift compensation, tremor averaging;
8). Analysis of eye motion trajectories and protection against eyelid blinking.

Comparing the general algorithms for HTS prototypes [20] and ETS shows that basic modules for image processing and computation are, in general, similar. Taking also in consideration that modular design of hardware allows studying the both systems with common equipment (differing, mainly, in the helmet module) the following description of the ETS algorithm and software prototype will include only specific software modules.

The difference in the algorithms consists, in different types of image models (eye – for ETS and head (RDU) – for HTS) and different equations for computation of coordinates. All algorithms are described, in [1, 2].

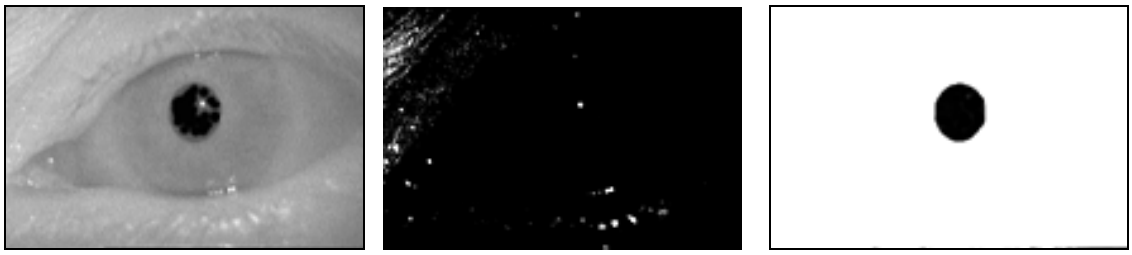
Examples of eye images taken by the ETS prototype during experimental testing are shown in Fig. 4.6. The results of the HTS and ETS experiments using the above mentioned SW modules are shown in [21].



a) left channel ETS b) right channel ETS

Fig. 4.6 ETS images of operator's eyes, four light spots are four corneal reflexes

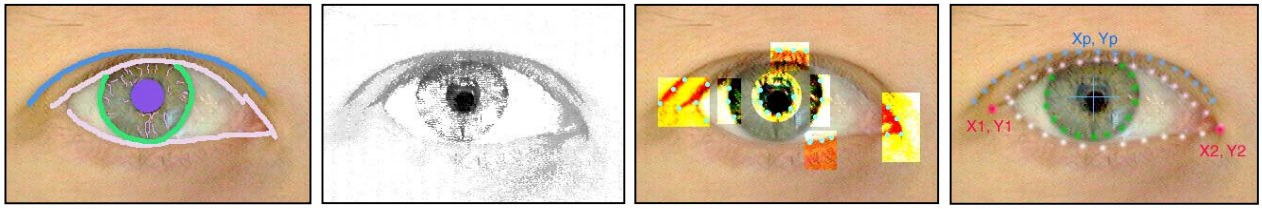
The image processing for active version of ETS prototype is illustrated in Fig. 4.7.



a) initial image, b) corneal reflex image c) pupil image

Fig. 4.7 Eye images obtained by experimental testing of active ETS prototype

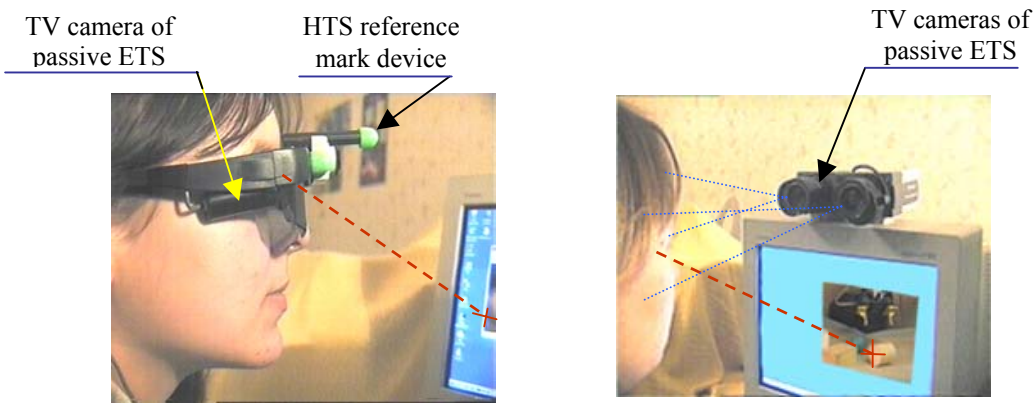
The image processing for passive version of ETS prototype (without IR LED's illumination) is illustrated in Fig. 4.8. Eye zone selection for eye image tracking in operator's real work is accomplished using additional information from the passive HTS, especially when losing the eye by involuntary closing or covering by hand.



a) application of the eye model b) selection of the pupil c) overlay of strobes d) processing results

Fig. 4.8 Image processing algorithm for passive ETS prototype

As an example of look direction measurement results utilization a development may serve of our proposals on 3D virtual cursor [22]. There is used a spatial cursor image in the prototype of passive ETS for directing operator's look at an object a robot is to interact with, (see Fig. 4.9).



a) helmet-mounted module of the ETS prototype b) PC monitor module of the ETS prototype

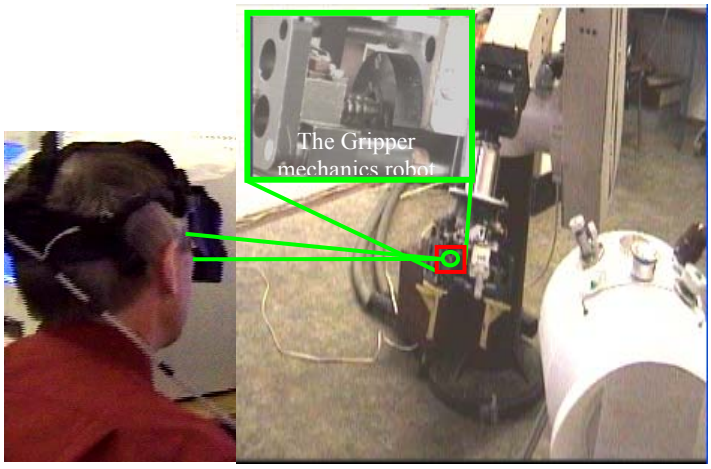
Fig. 4.9 Two variants of passive ETS prototype for the virtual cursor control

4.2.2. Controlling virtual cursor with ETS

One of the most promising applications of gaze pointing for robotics is developing the idea of the 3D "virtual cursor" (VC) used in systems of Augmented Reality (AR) and stereo vision. It provides high dynamic and natural easiness in VC pointing. The VC pointing by gaze may be realized only in combination ETS with a Head Tracking System (HTS).

Employing ETS in the telecontrol process may be very effective for dynamic acquisition of all needed data of an image fragment chosen by the operator's gaze.

The graphic means are being developed for helping operator in comprehension of a remote scene. It is possible, that the simplest way of data overlaying is a text giving information on a place of concern determined by a Line-of-Sight (LOS) position. For example, detailed technical information on earlier selected object added the current data, may be displayed (Fig. 4.10). The choice of a 3D virtual object's fragment by gaze with appearing tag of additional information characterizing this fragment is considered here.



a) man-operator b) displayed image of remote Work Zone image

Fig. 4.10 Using the «virtual cursor» with ETS control for obtaining detailed information

4.2.3. Analysis of ETS approaches to displaying a high resolution image

A video system for displaying a remote WZ image in the remote robot control technology is essential for significant expansion of operator's capabilities. While designing such system it is necessary to take in account physiological aspects of human perception of images.

An important feature of eye is a dependence of its resolution from image position on the retina and also from image velocity over the retina.

In the first instance, a cause is the design of a human retina optical system enabling a high definition only in the retina's center – the so-called fovea zone. The nearer to edge of retina the less is resolution of eye [23]. Man sees surrounding objects sharply by means of eye saccades from point to point and the whole scene is mentally integrated of sharp fragments.

In the second instance, at high speeds of objects eye cannot keep, there for an image becomes slurred. For speeds 30 °/s and 57 °/s the eye resolution goes down by 2 and 7 times respectively [24].

The abovementioned leads to a conclusion, that displaying a scene with high resolution the only a small zone of operator's concern is enough. All beyond its local zone may be displayed with worse resolution. It is enough to select a point looked at in a given moment and a small zone around it to create image with maximum resolution for man-operator.

But now it is important to provide a necessary degree of eye movement synchronism with movement of the high-resolution image zone (HRIZ). The time lag should not exceed the characteristic eye response time that is about 0,1s. This method of dynamic selection enables a considerable reduction of TV pass-band for image telecommunication without spoiling a perceived sharpness of picture.

Consider, for example, a system with a monitor having resolution 1024x1280 pixels and TV-channel with pass band equivalent to 300x400 pixels used and the workstation equipped with systems HTS and ETS. Co-operation of these two systems enables to determine a point on the monitor screen which the operator' look is focused in.

For example, the displaying system selects a 10% high-resolution image zone (HRIZ) with its center in looked-at point of monitor screen. Remaining 90 % of image is displayed with a lower resolution (by 2-3 times or more).

The HRIZ permanently following operator's look, in the result the operator perceives a high-resolution picture on the whole monitor screen. At the same time a TV pass band will be by 4-9 times less as compared with high-resolution image over all frame (screen). Thus an image with resolution 768x576 can be transmitted via a communication line with capacity less than needed for 300x300. The preliminary test results of ETS prototype with Virtual Cursor and HRIZ applications are presented in [21] and in Chapter 5 below.

4.3. Integration ETS& HTS with stereo Head Mounted Display

The proposed design scheme for the optical system makes possible combining in integral helmet-mounted module all indispensable elements of HMD, ETS and HTS prototypes.

By development of HMD prototype for research purposes of this Project the following considerations were take in account:

- the prototype should have minimal dimensions and weight;
- colour stereo image should be displayed at HMD;
- the see-through capability (image viewed on environmental background) should be provided;
- optical system should be possibly simple for wide FOV (no less than 30º) in horizontal plane;
- coupling with prototypes of ETS and optical HTS should be provided.

The HMD prototype comprises a binocular optical system consisting of two identical channels for left and right eye. Fig. 4.11 shows structural scheme of one optical channel coupled with elements of prototypes ETS & HTS and HMD.

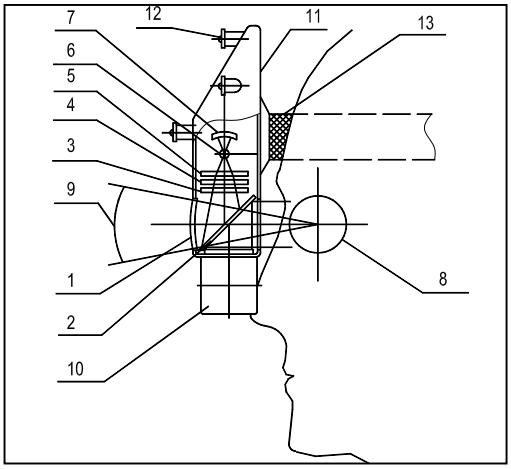


Fig. 4.11 Configuration of an integral helmet-mounted module (HM) with HMD, ETS and HTS

The optical channel is built according to an axisymmetrical reflector scheme and contains spherical semitransmitting mirror 1, flat semitransparent beamsplitter 2, translucent LCD 3, thermal filter 4, condenser 5, light source 5 and reflector 7. The LCD is placed in the focal plane of mirror 1. Light source 6 with condenser 5 and mirror 1 is optically concerned with eye rotation center 8.

Light rays from source 6 pass condenser 5, thermal filter 4 illuminate the operating field of LCD 3. Then, an image formed by the LCD are reflected from beam splitter 2, focused with mirror 1, pass beam splitter 2 again and come to user's eye 8 seeing the imaginary image at the optical infinity.

The focal length of mirror 1 is taken with the view that FOV 9 of the system were 30º x 22º, in horizontal and vertical planes, respectively. The center of curvature of the mirror is combined with eye center of rotation 8 for minimization of aberration.

Elements 10 of ETS prototype may be accomodated in the lower part of the helmet-mounted module and, by that, the eye-directed surface of beam splitter 2 should have a selective coating reflecting wavelengths of ETS.

The HM of HTS may be combined with HM by placing in the upper part of HM housing 11 of necessary number of IR LEDs 12 (no less than 4) belonging to HTS. The design of HMD (Cy-visor) is used in the integral HM see in Fig. 4.12.

Fig. 4.12 Design of Cy-visor DH-4400 HMD

Summary (for Chapter 4)

Our research is aimed at obtaining required accuracy and reliability of eye positioning (direction of look) using optic methods. In the process of research and development of the ETS prototype the following tasks were fulfilled:

1). A comparative analysis was made and criteria were formulated for characteristics and advantages of the optical principle as compared with other principles of ETS realization.
2). Optical and TV methods were developed for protection against interference of external illumination that ETS may suffer from at the control post.
3). Several optical and algorithmic methods were proposed for adjustment and relating ETS measurements to the head coordinate system and the system of controlled objects (robots).
4). Different methods for image processing were developed and realized (filtering, selection, identification and image motion analysis), ones using colour information in their number.
5). Study was accomplished of perspective methods for structuring information contained in the real images using 3D models of human visual system and, also, using analysis of eye look trajectories.
6). Experimental study of operator's sensory-motor functions by robot telecontrol, and as such, methods had been studied for coordination of look and head motion and, also, coordination of look and hand motion.
7). Methods had been studied for the synergy of tactile, force-torque and visual aspects by manual telecontrol operations including that using eye tracking systems (HTS, ETS&HMD).

Gaze control applications are the following:

A). Virtual cursor (VC) displayed on a stereo (3D) monitor;
B). High resolution image zone (HRIZ) on computer monitors, displays for collective use and helmet-mounted ones;
C). Zone of higher interest displaying more details or additional information while viewing computer virtual images combined with actual images;
D). Stereo cameras (control of convergence) tracking an actual object;
F). Passing point-of-attention data to any other operator (remote LOS exchange).

Chapter 5. Experimental research of a novel man machine interface for telerobotic using an eye tracking system

Some experimental results of a Man-Machine Interface (MMI) development for the robot telecontrol basing on a system tracking man-operators' gaze are presented.

In the framework of this task an Eye Tracking System (ETS) prototype is studied as a part of a functionally multilevel dynamic system controlled by eye and head motions.

The ETS had been tested in the Hardware & Software Complex (HSC) facility using the following general criteria for expert assessment:
- Effective look pointing while controlling the virtual cursor (VC) with ETS;
- Possibility of control with ETS of the high-resolution image zone (HRIZ);
- Simplicity of training by showing a natural gaze motion.

5.1. Experiments with ETS controlling the virtual cursor

The basic task of the experimental study is to inquire into principles and mechanism lying in the base of operator's sensory-motor functions for robot telecontrol. As an example of results in eye tracking with the ETS prototype development may serve of 3D virtual cursor (VC). In the ETS prototype VC is used for showing a robot what object is to manipulate with.

The VC can be used for measurement of distances and shapes of 3D image of real scene objects similar to that as it is made in CAD-systems (virtual meter). It is possible to contour boundaries of virtual or real objects, to set section planes or to point a path of the robot gripper's movement.

Using ETS & Head Tracking System (HTS) for control of camera movement enables more detailed scanning of scene by natural motions of head. Another mode is observation of manipulation objects in the work scene. A video record resulting from survey of Work Zone (WZ) or observation of objects is further transformed into a 3D panoramic picture.

While operating with a wide size picture man needs not see all picture with the same resolution. We have experimentally studied possibility of a High-Resolution Image Zone, (HRIZ) selected with ETS.

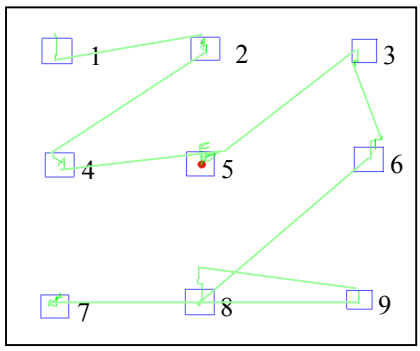
A work is begun for teaching a robot-manipulator in trajectories by way of showing. The operator observes WZ with ETS using his own experience of it. Robot repeats the procedure complying with an order and trajectory of man's LOS. And prior to an action robot control system may move the cursor over the obtained panorama image, vary scale of fragments and, upon ascertaining that all is right, executes the actual movement.

The gaze control using ETS & HTS enables solving a task of teaching intelligent robot control systems in WZ observation, for example, for performing assembly operation's in outer space.

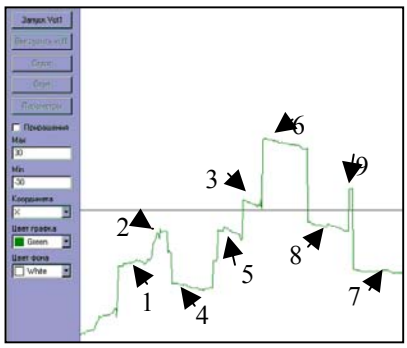
Experimental study of control processes for scene observation and choosing objects for robot gripping with the help of ETS using the 3D virtual cursor (VC) had been carried. An additional use of VC for pointing with ETS enables, in our case, a predictive control with head or hand. The PC monitor displays work scene objects tagged with coordinate data and the cursor moving under control of ETS.

The ETS signal is outputted to Robot Control System (RCS) as a signal for targeting a possible movement of the robot camera. For confirming the target positions pointed with eye various kinds of confirmation signals are used: voice, gesture, pressing a key. These are the commands for executing the predictive control already shown by gaze.

On this stage of work we have accomplished experiments for testing the virtual cursor control while pointing with ETS to square areas in the test picture. The path of gaze fixed on targets (1…9) for about 20 s is shown in Fig. 5.1 a.



a) results of the cursor pointing



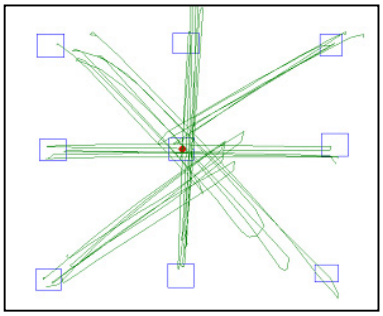
b) diagram of eye movement

Fig. 5.1 Trajectories of look while stabilization look at target areas

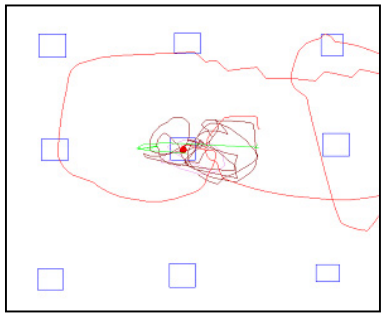
The experiments showed a possibility of look stabilization with acceptable accuracy. The small vertical drift of gaze is caused by operator's breathing and therefore unstabilized head positions (see diagram Fig. 5.1 b).

Results of experiments presented in Fig. 5.2-a showed a high accuracy of repeated look's hitting a target zone of a test picture. Fig. 5.2-b shows VC movement while operator's keeping look in a definite point of the test picture and in the same time turning his head.

This experiment confirmed the fact that while operating with ETS operator's head movements should be corrected for with the help of HTS. By the virtual cursor control with ETS unwanted movements of head must be compensated for with HTS.



a). VC movement at saccade eye movements from center to periferal points with head stabilized

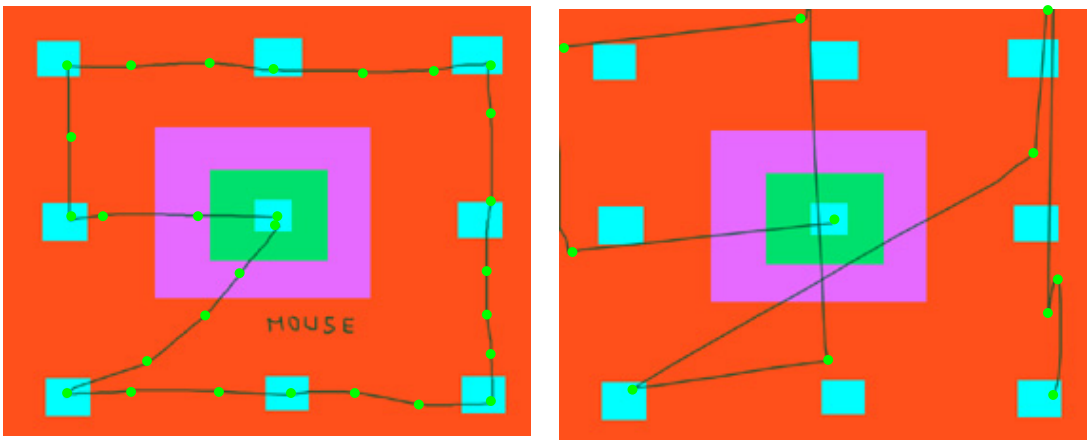


b). VC movements while look is stabilized at the centre and head smoothly moves

Fig. 5.2 The VC movements controlled by gaze with (or no) head stabilization

Prior to work the ETS system must be adjusted for an individual operator, its alignment and choice of initial parameters. The purpose of this adjustment is to enable the operator to see the whole monitor screen with natural movements of eyes.

With eye pointing a command for execution may be given either by pressing a button (requiring both the cognitive and motor time) or automatically upon some time of look fixation. The experiments showed the advantage of ETS control of the virtual cursor (as compared with conventional "mouse" and joystick) in speed and hitting at one attempt (Fig. 5.3). In our experiment VC was controlled with mouse and the ETS prototype. The experiments proved that ETS is two times faster than "mouse". Moreover, cognitive time prior to start is for ETS (50-100 ms) less than that for "mouse" (200-300 ms). With a good calibration, as the experiments showed, VC pointing with ETS takes time 30-40 % less than that of "mouse" and is done in the most natural way.



a) path of pointing by “mouse” b) path of pointing by gaze (ETS)

Fig. 5.3 Pointing by “mouse and by gaze (green point show 50 ms time-mark)

5.2. Experimental study of the effect of presence in WZ using panoramas controlled with ETS

5.2.1. Image acquisition from a single point of shooting

In the simplest case a number of WZ images are obtained from a rotating camera around some optical center. “Mosaicing” (integration fragments into a single picture) known in cartography and processing aero and space photographs employed for robots-manipulators gets further development.

The basic difference consists in the active and controlled character of video shooting when a point of view and camera orientation in space are set by a robot. Besides that, work scene images are to be processed in real time mode (or near to) for prompt imaging 3D scene for telerobotic control purposes.

Most often, a need of composing a panorama is imposed by a requirement of high resolution, which cannot be obtained with one photo. The mosaic image in this project is produced for imaging real work scene while operator controls a remote robot-manipulator.

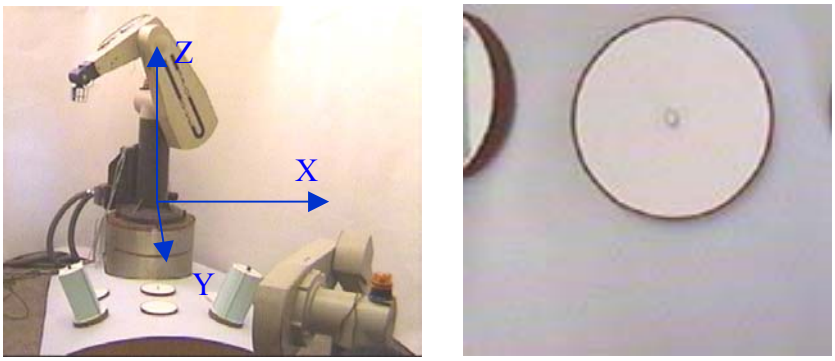
When a single-center panorama is used the resulting panorama should have cylindrical or spherical form for preserving natural spatial position of scene objects corrected for projective distortion. Therefore, for creating panorama one needs to know camera orientation for each fragment of it.

This orientation may be simply obtained from the robot control system ensuring enough high accuracy of camera positioning by the gripper. The camera optical center is obtained by means of some calibration method, as one a method using a number or reference points on real environment [25].

When the camera orientation is not known with required accuracy, e.g. while using an elastic-link manipulator or by arbitrary camera position on the gripper, one uses geometric proportion between image fragments [26]. An image field of view is generally much narrower than field of view of man – that is why a mosaic is built from a number of image fragments (see Fig.5.4).

The process of synthesis unites a sequence of fragments using a set of transforms and eliminates overlaps. The correlation method may have a slow convergence and often needs initial hand splicing of fragments. To overcome these disadvantages the characteristic feature selection is used 11].

The characteristic feature analysis reduces the mass of computation and proves operability at high turn angles and with changing image scale. All procedures may be accomplished automatically [26].



a) work scene of robot b) fragment of the work scene from CCD

Fig. 5.4 Work scene survey and observation of objects with robot-borne camera

5.2.2. Algorithm for synthesis and displaying realistic 3D images using ETS & HTS

One of ways for video data representation is a creation of panorama images as the base for composing GM of scene. Such presentation of data has a number of significant advantages:

1). Maximal information volume.
For cylindrical panoramas the vertical field-of-view (FOV) dimension is generally up to 100°. An image is virtually seen by operator as being projected onto the cylinder, enclosing operator. The operator may turn his head to right and left for surveying it, the same being equipped with ETS & HTS. For spherical panoramas the vertical FOV dimension may reach 180° and the operator may turn his head in 3D.

2). Minimal memory available.
As compared with a set of photographs or video records, the same data being compressed into a panorama without any loss at 100 times.

3). Arbitrary choice of an angle range and scale.
As compared with a video, when the operator sees the video by "eyes of a camera-man", with usage of panoramas he himself chooses place and look duration of such or other fragment making it by a natural motion of head (gaze) or hand. The arbitrary and dynamic choice of angle range and scale is possible now while surveying the work scene, the same with ETS & HTS, what enables the most easy way for realizing interactive choice a video data (MMI for video streams).

5.2.3. Work scene survey and observation of objects

The first algorithm's mode using Robot Remote Viewing (RRV) is a preliminary work scene survey. In this case we use work scene's mock-ups having limited dimensions (see above Fig. 5.4a).

The human-operator's eyes view (in real telecontrol practice) simultaneously only a little fragment of the work scene on the monitor, that it is significantly spoiled his orientation in space (see above Fig. 5.4b).

Therefore, it is desirable prior to actual work to give him the most full knowledge of work scene, e.g. with the help of panorama pictures composed of earlier obtained video data.

A wide-field camera with 92º field of view was used for making a video record (format *.avi**,** about 300 MB) of the maximal space. Then, using a special SW, a file in *.bmp format (4MB) was created with distortions cleared. It is a panorama showing the whole room. But such a picture is not easy to look at as a whole (Fig. 5.5).

Fig. 5.5 Dynamic control of panoramic image with ETS

But using a special SW this file may be surveyed in a panorama mode controlled with the “mouse” or ETS & HTS choosing a needed fragment of it (see Fig. 5.6).

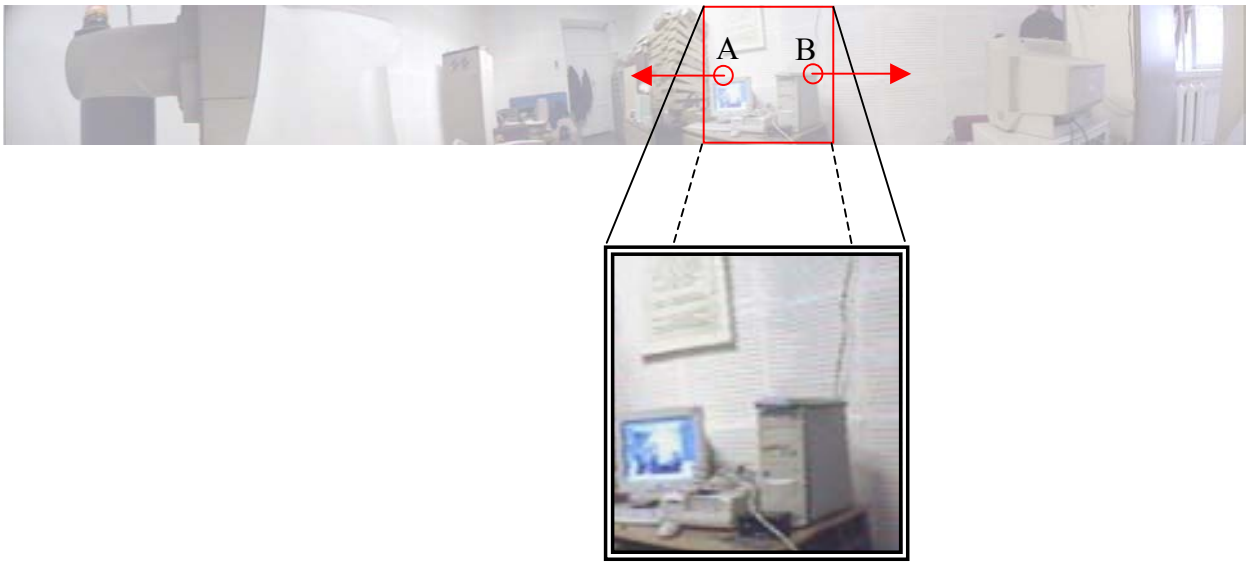


Fig. 5.6 Dynamic control of panoramic image of the CP room controlled by ETS & HTS

While surveying an obtained panorama using ETS & HTS it is important that sequence of pictures in the screen window were changed not with a random, chaotic turning of eyes but only being commanded with looks fixing some time at the boundary of adjoining frames.

Multiattitude integral pictures were created for observation of objects providing information of objects from various points of view (Fig. 5.7).

Fig. 5.7 Multipoint-of-view integral pictures

Because, while observing an object, relative positions of objects and background change one cannot obtain a unified coherent picture of the whole surface of object. To have a full knowledge of object one needs to use a multiaspect picture with a possible immediate access to an aspect of interest. This part of studies will continued in next stage of researches.

During experimental studying panorama pictures with the help of ETS & HTS the following peculiar features were revealed:

1) A panorama picture needs a compact size of file to record as compared with the initial set of photographs or video fragments yet keeping full video data. The experimental study showed a 100 times compress file size while preserving a required quality of picture.
2) While watching a panorama user does not see all by eyes of the camera-men, he himself chooses a direction and time of observation of any fragment commanding it by a natural movement of hand or head (glance). Now it is possible to choose at one's will an aspect and scale, e. g. with the help of ETS & HTS, enabling the most easy way of operation with video data.
3) When video data, obtained both by operator himself and cameras in the work scene, are not available in full, the picture may be supplied with fragments of a computer model. In this case, known as Augmented Reality (AR), a precise automatic registration of virtual objects with real ones is a main problem this Project deals with [17].
4) An actual field of view provided by a camera depends from many factors: lens design, optic parameters, FPA size etc. Therefore prior to work the parameters are to be defined and the camera must be calibrated.
5) Using a HTS for observing panorama pictures it's scanning should be controlled not only with turning head but with speed of it turn too.
6) During the experiments it was found that in making a panorama the camera is not to be turned too fast. It must be done smoothly and with such speed as enable man to see an object in all details. A turn at 90° should take at least 8 seconds.

The technical demands to accuracy of registration of real and model images in AR system are imposed by basic physiological individualities of man-operator. In the process of perception of spatial properties of object the human eye very finely perceives difference in shapes of objects and their parts (form recognition threshold) and, also, comparative sizes of objects (size recognition threshold).

Curvature of lines is perceived when it has angular measure of 9º, straight line bent is perceived at a less angle (about 5-8º). Thus, the eye may imperceptibly diverge from the straight line at 5-9º. A shift off a straight line is perceived having alike measure. It is interesting to note that a recess is more recognizable than a bump.

While comparing two rectangulars with equal bases we recognize difference in their height only 1/50 – 1/60 of it. The eye finely feels change of angle, 40º recognition threshold is near 47' [27]. But the perception is individual depending from experience, skill, professional interest and specific task. That is, a professional will better perceive important details and skip the secondary ones.

Every act of perception should be considered as a process and not a finished fact. One recognizes two stages of perception: momentary vision and detail observation. In the first moment one recognizes object or catches some characteristic property. So, passing a street we catch by eye an approaching car and at once make distance to it. Knowledge on objects and phenomena obtained in such way helps man to act in dynamic situations. For more full and profound knowledge man needs the detailed observation what is not always possible in real work of operator.

For providing capability of simultaneous vision of detailed images and entire picture we have proposed a mode of panorama synthesis continuously updated with current frames. In other words, synthesis and updating in the real time mode. Fig. 5.8 shows a procedure augmenting a real image with virtual image of SS model (augmented reality– AR).

An opposite procedure is possible when a virtual image is augmented with real objects or texture (Augmented Virtuality – AV) [26].

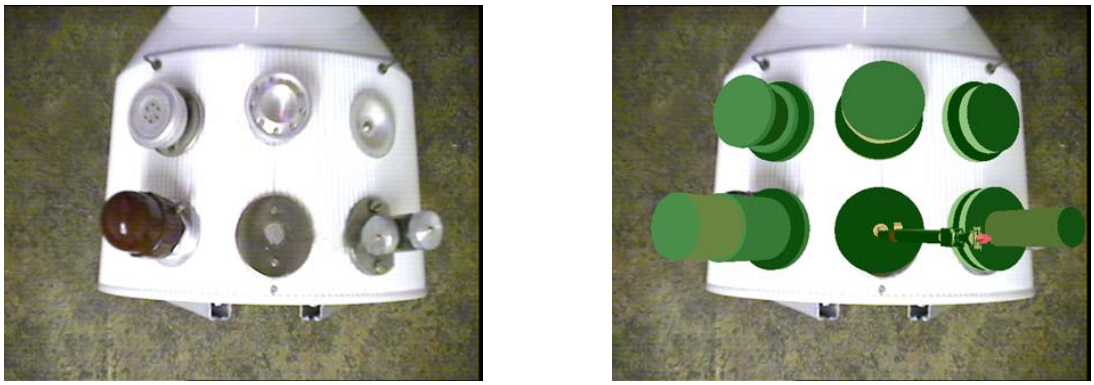


Fig. 5.8 The augmenting a real with virtual image

In the latter instance the panorama synthesis, currently realized as overlaying fragments of texture on virtual image (3D model), and also requires accuracy in registration of edges, scales aspects of real fragments and virtual images.

It is known that the kinesthetic feeling has a significant role in forming any types of perception. The kinesthetic feeling combined with other simultaneous perceptions forms a functional system of perception vividly conveying phenomena and objects of real life.

The vision of man shows itself in two forms: visual-motoric and visual-kinesthetic. The first consists, essentially, in movements of eyes and head. Owing to its perfect motoric system the eye can perform high variety of motions.

These motions, as a rule, define contours, curvature of shapes, change of direction and in contours spatial surfaces and other specific features of visually cognizable objects. Man obtains a sequence of visual impressions from different parts of objects intervened with turns of eyes and head, contraction of eye or other muscles followed by kinesthetic feeling.

Eye's rotation immediately sends to mind information on change in position of point currently looked at relative to one that was fixed before. Thus, the process of visual perception cannot be separated from the process of thinking.

Another form shows itself in object perception both with sight and touching with hand. The close interaction of visual and tactile perception of real life is seen in many kinds of human activity both in cognition and practical work. The more the skill the less the visual control and many motions begin to be primarily controlled by the tactile-kinesthetic system [28].

Therefore, it is integrating a whole picture out of fragments in definite sequence enables fidelity of perception even when some data are lost.

5.2.4. Creating a mosaic picture with missing video data

With missing video data a panorama may be made by a method of registration and mosaicing [22]. The method provides registration of actual fragments of objects with geometric model ones. Overlaying GM fragments is done taking in account points of view they are seen from what significantly betters the perception.

A sequence of fragments in space should comply with a natural order of objects or observation of scene by man. That is convenient and economical in time for searching a needed fragment. The choice of fragment displayed on the monitor is done with ETS (VC and HRIZ).

Fig. 5.9 shows spatially disordered fragments of actual images of a work scene (SS Mock-up) and Fig. 5.10 shows them spatially organized and supplied with GM fragments.

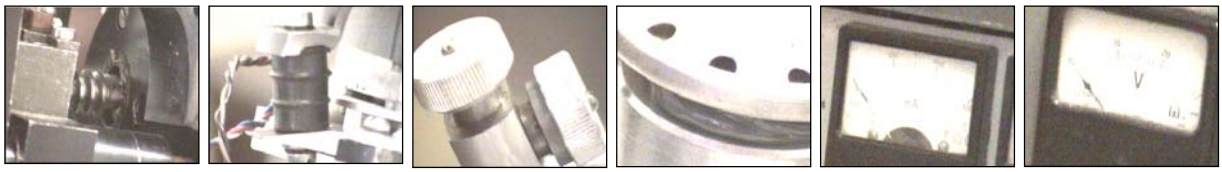


Fig. 5.9 Spatially disordered fragments of actual images of the SS' Mock-up

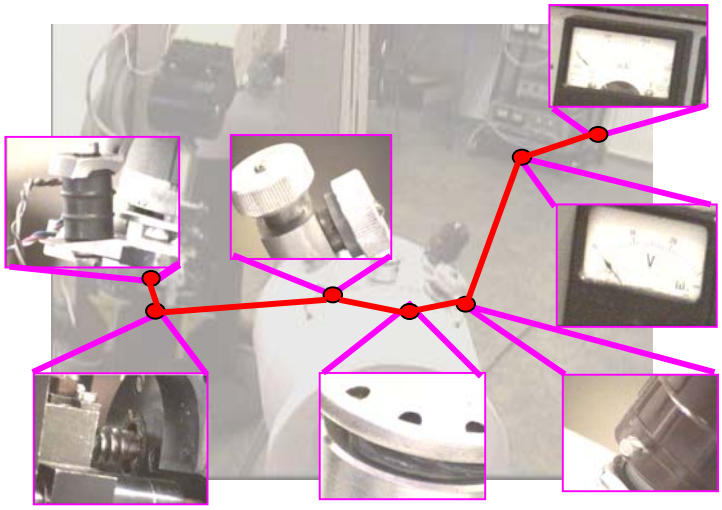


Fig. 5.10 Spatially organized actual fragments of the SS Mock-up presented in a form of graphic video model

A sequence of other fragments of Space Station Mock-up images presents the results of a scene observation (see Fig. 5.11). That is not convenient and demand long time for searching a needed fragment. The choice of necessary fragment with ETS by natural way is done simpler.

Fig. 5.11 shows the next example of frames SS Mock-up and Fig. 5.12 shows spatially organized ones.

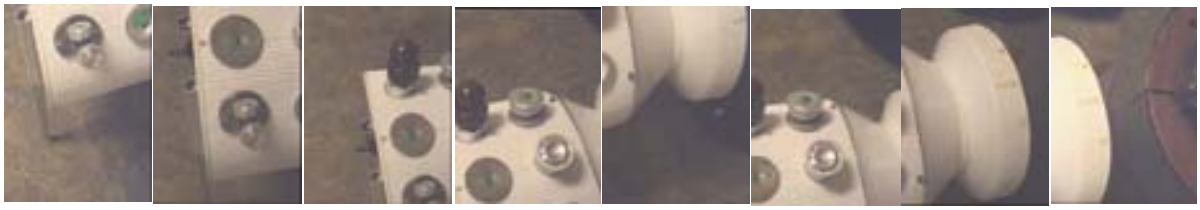


Fig. 5.11 Spatially disordered frames of actual images of the SS Mock-up (original *.avi file =17,9 Mb, single frame *.bmp = 225 Kb)

Fig. 5.12 Spatially organized frames of the SS Mock-up, (*.bmp file 1,076 Mb)

5.3. Preliminary experimental study of work scene representation with the high resolution image zone (HRIZ) controlled with ETS

In Chapter 4 was shown, that displaying a scene, with high resolution at small zone of operator's concern is enough. All beyond its small zone may be displayed with worse resolution. It is enough to select a point looked at in a given moment and a small zone around it to create image with maximum resolution [23].

But now it is important to provide a necessary degree of eye movement synchronism with the movement of the high-resolution image zone (HRIZ). The time lag should not exceed the characteristic eye response time that is about 0,1s.

This method of dynamic selection enables a considerable reduction of TV pass-band without spoiling of perceived sharpness of picture. To realize this method one needs equipment enabling the permanent tracking of gaze direction. As such equipment the systems HTS and ETS can be used. Now, if the window of interest is 10% of full picture the file size is 100 times less than that of initial image. The initial image (15 Mb) using format JPEG 2000 we reduce to the file size 143 Kb.

If, then, the compressed picture is overlaid with a sharp window of interest (Fig.5.13) the resulting file size (~ 200 Kb) is still far less yet preserving a high definition detail of interest. This single zone of high definition changes its position on the slurred picture as look moving over it.

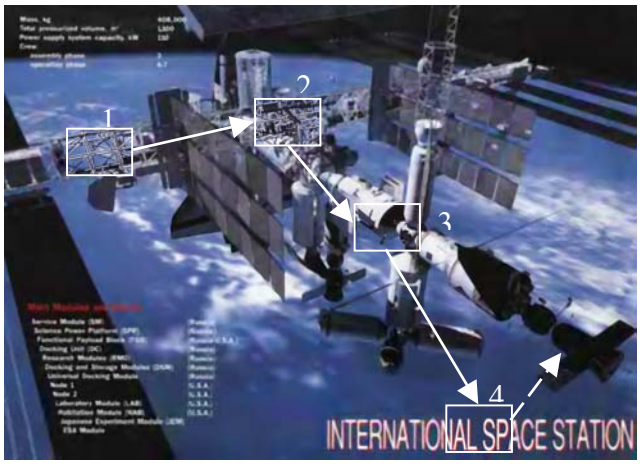


Fig. 5.13 The sequence of a high resolution zone movement by ETS control

A zone of interest is chosen with ETS. It follows movement of look. Fig. 5.14 (below) shows consecutive steps in examining the picture of Space Station's (SS) Mock-up.

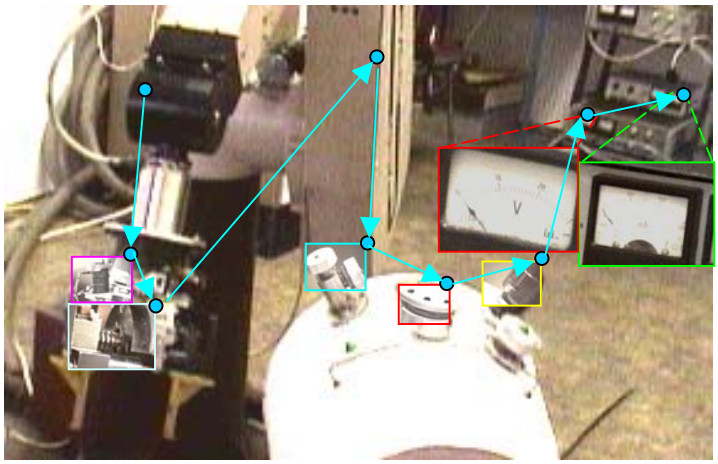


Fig. 5.14 The sequence of high resolution zones on SS' Mock-up image moved with ETS, some HRIZ' image fragments are zoomed to

A zone of interest is chosen with ETS too. It follows movement of man-operator gaze during the telecontrol process.

The experiments showed a possibility of look stabilization with acceptable accuracy (more less than 15 angle min). The small vertical drift of gaze is caused by operator's breath. The some results of experiments with ETS are shown in Appendix 2.

Summary (for Chapter 5)

The following functions have been considered as tasks for the gaze control:

1). 6D control of the virtual cursor (VC) displayed on a stereo (3D) monitor;
2). 3D control of the high-resolution zone in displayed image (HRIZ), for helmet-mounted displays, computer monitors and any displays for collective use;
3). 3D control of the image zone of concern (with added image details and augmented information) while looking computer virtual images combined with actual images;
4). 3D control of the camera pair (control of convergence) while tracking images or actual objects;
5). Recognition of gesture, mimic and articulate commands for realization of the intelligent MMI;
6). Passing look-of-sight data to other operator ("remote look exchange").

Conclusion

Main results of activities on Task 5 of Project # 1992p:
1). Advanced methods of MMI based on HTS&ETS have been developed for robot-manipulator telecontrol and for telecontrol of other spatially moving objects.
2). HTS, HTS+ and ETS prototypes were fabricated for robot telecontrol MMI.
3). The Hardware & Software Complex (HSC) facility with virtual models and real robot-manipulators "PUMA" was designed for HTS, HTS+ and ETS prototypes verification and testing.
4). Experiments and tests of HTS, HTS+ and ETS prototypes were fulfilled and main technical parameters were verified successfully with robot-manipulators of HSC.

Main methods used for development of the intelligent MMI are the next:
1). Recognition of head/hand images on real background using 3D frame-structural models;
2). Measurement of spatial position and orientation of head/hand in real time mode;
3). Coordination of head movement and control of virtual and real images;
4). Teaching MMI by man-operator's showing the characteristic motions of head and hands.

The Task 5 of the Project is aimed at obtaining required accuracy and reliability of eye position measurements using optic methods. In the process of research and development of the ETS prototype the following tasks were fulfilled:
1). A comparative analyses had been made and criteria had been formulated for optical methods as compared with other ones for ETS realization.
2). Optical and TV methods have been developed for protection against interfering outer illumination, which ETS may suffer from at the control post.
3). Several optical and algorithmic methods were proposed for adjustment and tying ETS measurements with the head coordinate system and the system of controlled objects (robots).
4). Different methods for image processing have been developed and realized (filtering, selection, identification and image motion analysis), ones using colour information in their number.
5). A study has been accomplished of perspective methods for structuring information contained in real images using 3D frame-structural models.
6). An experimental study is under way of operator sensor-motor functions in robot telecontrol, and as such, methods are being studied for coordination of look and head motions and, also, coordination of look and hand motions.
7). Methods are being studied for additional use of the tactile and force-torque aspects for robot telecontrol using HTS, HTS+ and ETS.

The operation of HTS and HTS+ with the space station mock-up and actual robot-manipulators is realized. Experimental studies of algorithms and SW for telecontrol are carried out:
1). The position and speed control is ascertained for commanding linear and angular displacements with natural movements of head and hand.
2). Limits are established of the zone of stable control for linear and angular movements of RDU. Large ranges of hand movements are demonstrated and necessity is established of hand position fixation for ease of control.
3). Experiments are accomplished for estimating control process dynamics in various operation modes.
4). The efficacy of passive HTS and HTS+ is tested while operating with color reference marks on actual background of the control post (CP), with sunlight reflected from walls and with lamps appearing in the camera's FOV.

Some applications of MMI system based on HTS (HTS+) prototypes:
- Advanced computer interface for gesture exchange with PC;
- Pseudo-holographic effect of perception 3D images;
- MMI for home (office) servicing robot-like devices;
- MMI for telemedicine systems, medical robots, etc;
- Simulators of real time control process (nuclear station, aviation, and others);
- Remote control of the observing camera (DOOMe, Web-cameras, security ets);
- MMI for telerobotic control with effect-of-present in remote WZ;
- MMI for multi-robotic systems (multi-channel control by usage two hands + head).

The following applications have been studied for gaze control:
1). 6D control of the virtual cursor (VC) displayed on a stereo monitor;
2). 3D control of the high-resolution zone in displayed image (HRIZ) for helmet-mounted displays, computer monitors and any displays for collective use;
3). 3D control of image zone of perception concern (with added image, details and augmented information) while viewing computer virtual images combined with actual images;
4). 3D control of the mobile camera pair while tracking images or actual objects;
5). Recognition of mimic and articulate commands for the intelligent MMI;
6). Passing point-of-gaze attention data to other operator ("remote look exchange").

References

1. Interim Report #3 "Eye Tracking and Head-Mounted Display/Tracking Computer Systems for the Remote Control of Robots and Manipulators". Project #1992p, Task 5, May 2002.
2. Interim Report #4 "Eye Tracking and Head-Mounted Display/Tracking Computer Systems for the Remote Control of Robots and Manipulators". Project #1992p, Task 5, Nov. 2002.
3. Workshop Conference, Binghminton, March 2002.
4. Interim Report #4 "Technology for the Creation of Virtual objects in the Real Word". Project 1992p, Task 6, November 2002.
5. Bunjakov B.A., Burdygin A.I., Kolesnik A.M., Nechaev A.I., Chernakova S.E. "Algorithm of automatic guidance of the autonomous robot". Proc. of X technological conference "Extreme robotics". St.-Petersburg 1999.
6. A. I. Burdygin, F. M. Kulakov, A. I. Nechaev, S. E. Chernakova: “A multiphase method and algorithm of measurement the spatial coordinates of objects for teaching of assembly robots”, SPIIRAS Proseeding, Issue No.2. - SPb: SPIIRAS, 2001.
7. F. M. Kulakov, A. I. Nechaev, S. E. Chernakova: “Modeling of Enviroment for the Teaching by Shoving Process” the Proceedings of SPIIRAS, Russia, St-Peterburg, 2001.
8. N.Lauinger: “Diffractiuve 3D grating-optical image processing: an interference-optical Volterra filter resonator”, Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, Proceedings of SPIE (2001) (p.p. 61-69).
9. Interim Report #1 "Technology for the Creation of Virtual objects in the Real Word". Project 1992p, Task 6, May 2001.
10. G.A. Watson, T.R. Rice: “Sensor bias estimation and compensation for improved track correlation”, Acquisition, Tracking, and Pointing XV, Proc. of SPIE (2001) (p.p. 112-125).
11. Interim Report #5 "Eye Tracking and Head-Mounted Display/Tracking Computer Systems for the Remote Control of Robots and Manipulators". Project #1992p, Task 5, May 2003.
12. Interim Report #2 "Eye-Tracking and Head-Mounted Display/Tracking Computer Systems for the Remote Control of Robots and Manipulators". Project #1992p, Task 5, Nov. 2001.
13. F. M. Kulakov, A. I. Nechaev, A.I. Efros, S. E. Chernakova: “Experimental study of man-mashine interface implementing tracking systems of man-operator motions” the Proceedings of Sixth International Seminar on Science and Computing, Moscow, Russia, September 2003.
14. F. M. Kulakov, A. I. Nechaev, A.I. Efros, S. E. Chernakova: “Hard & software means of man-machine interface for telerobotic using systems tracking man-operator motion” the Proceedings of Sixth International Seminar on Science and Computing, Moscow, Russia, September 2003.
15. V.P. Bogomolov, S.I. Kostin “Robotic operation in space” Space magazine, May-June 1997.
16. F. M. Kulakov, A. I. Nechaev, A. I. Efros, S. E. Chernakova: “Hard & software means of MMI for telerobotics using systems tracking human-operator motions”, Proc. of III International conference «Cybernetics and technology of XXI century» October, 2002, Voronezh, Russia.
17. Interim Report #5 "Technology for the Creation of Virtual objects in the Real Word". Project 1992p, Task 6, May 2003.
18. S. Grange, F. Conti, P. Rouiller, C. Baur “The delta Haptic Device as a nanomanipulator” Microrobotics and Microassembly III, Proc. of SPIE 2001.
19. Interim Report #1 "Eye-Tracking and Head-Mounted Display/Tracking Computer Systems for the Remote Control of Robots and Manipulators". Project #1992p, Task 5, May 2001.
20. F. M. Kulakov, A. I. Nechaev, A.I. Efros, S. E. Chernakova: “Novel man-machine interface for telerobotics using eye tracking systems” the Proceedings of Sixth International Seminar on Science and Computing, Moscow, Russia, September 2003.
21. F. M. Kulakov, A. I. Nechaev, A.I. Efros, S. E. Chernakova: “Experimental research of novel man machine interface for telerobotics using eye tracking systems” the Proceedings of Sixth International Seminar on Science and Computing, Moscow, Russia, September 2003.
22. S. Harasaki and H. Saito: “Vision based overlay of a virtual object into real scene for designing room interior”, Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, Proceedings of SPIE (2001) (p.p. 545-555).

23. Kravkov S.V. Eye and its operation, M, Leningrad, edition SU Academy of Sciences, 1950.
24. Koroleonok K.N. On cinematographic perception / Problems of general psychopathology. Sciences works collection. Irkutsk, 1946, pp. 198-214.
25. Henrik Haggren, Petteri Pontinen, Jyrki Mononen “Cocentric image capture for photogrammetric triangulation and mapping and for panoramic visualization” Part. of the IS&T/SPIEW Conference on Videometrics VI, (1999).
26. Work Materials #4 "Technology for the Creation of Virtual objects in the Real Word". Project 1992p, Task 6, November 2002.
27. Andreeva E.A., Vergiles N.U., Lomov B.F. "The mechanism of eye elementary motions as a tracking system" in book: Motoric components of vision" M, 1975.
28. Iarbus A.L. "The role of eye movement in the process of vision" M, 1965.

ACRONYM	DEFINITION
2D (3D)	Two Dimension (Three Dimension image or model)
6D	Six Dimension (coordinates)
AI	Artificial Intellect
AV	Augmented Virtuality
AR	Augmented Reality (technology)
CAD	Computer Artificial Designed
CCD	Charge Coupled Device
CCU	Camera Control Unit
CP	Control Post
CU	Camera Unit
DOF	Degree Of Freedom
ETS	Eye Tracking System
FOV	Field Of View
FPA	Focal Plane Array
FSM	Frame-Structural Model
GM	Geometrical Model
HM	Helmet Module
HMB	Head (hand) Motion Box
HMD	Helmet Mounted Display
HRIZ	High Resolution Image Zone
HSC	Hard & Software Complex
HTS	Head Tracking System
HTS+	Hand Tracking System
IR	Infra Red
IR LED	Infra Red Light Emission Diode
IS	Image Signal
JPEG2000	Joint Picture Engineering Group 2000 standard
LED	Light Emission Diodes
LCD	Liquid Crystal Display
LOS	Line Of Sight
MMI	Man-Machine Interface
MPEG	Motion Picture Engineering Group standard
OS	Optical Signal
PAL	TV colour standard
PC	Personal Computer
PCI	PC slot standard Interface
RC	Remote Control
RCS	Robot Control System
RDU	Reference Device Unit
RMS	Root Mean Square
RRC	Robot Remote Control
RRV	Robot Remote Viewer
RV	Remote Viewing
RVS	Robot Vision System
SS	Space Station
SW	SoftWare
TV	Tele Vision (system, signal, camera)
TMS	Television measurement system
UAV	Unmanned Aviation Vehicle
USB	Universal Serial Bus (PC interface)
UWV	Underwater Vehicle
VC	Virtual Cursor
VPU	Video Processor Unit
VS	Video Signal
WM	Work Materials of Report
WZ	Work Zone (of remote robot-manipulator)

Appendix 1

Results of experimental testing of the HTS and HTS+ prototypes

The experiments were carried out using 4-mark reference devices for HTS and 3-mark reference devices for HTS +, shown above in materials of this Report Chapter 2. The preliminary results of experiments are presented below.

1. Estimate of RMS coordinate measurement error

1.1. Active HTS prototype, 4-mark RDU, distance 700 mm

The output data of HTS are presented (without filtering) on diagrams Fig. A1.1-A1.5, a full size of diagram: 0.2 pixels for Y and 100 sec. for X.

Conversion of linear measure to that in pixel at distance 700 mm corresponds to: σ_x, σ_y = 20 μm ≅ 0,006 pixels

The results of experiments with active HTS prototype, RMS error see in Table A1.1.

Table A1.1

It	Name of parameter	Units	X	Y	Z	φ_x	φ_y	φ_z	Note
1.	RMS (σ) error (without filtering)	μm, arc. min.	23	33	33	1	2,2	2,6	4-mark RDU HTS
2.	RMS (σ) error (with median filtering)	μm, arc. min.	18	28	40	1	1,8	1,8	

1.2. Active HTS+ prototype, 3-mark RDU, distance 700 mm

The results of experiments with active HTS+ prototype, RMS error see Table A1.2.

Table A1.2

it	Name of parameter	Units	X	Y	Z	φ_x	φ_y	φ_z	Note
1.	RMS (σ) error (without filtering)	μm, arc. min.	30	43	52	1,2	2,8	2,9	3-mark RDU HTS+
2.	RMS (σ) error (with median filtering)	μm, arc. min.	22	32	64	1,0	2,0	2,0	

On diagrams, presented below in Fig. A1.1, value σ_x = 0,009 pix corresponds to 30 μm, as related to the full diagram size ΔX_G = 106 mm.

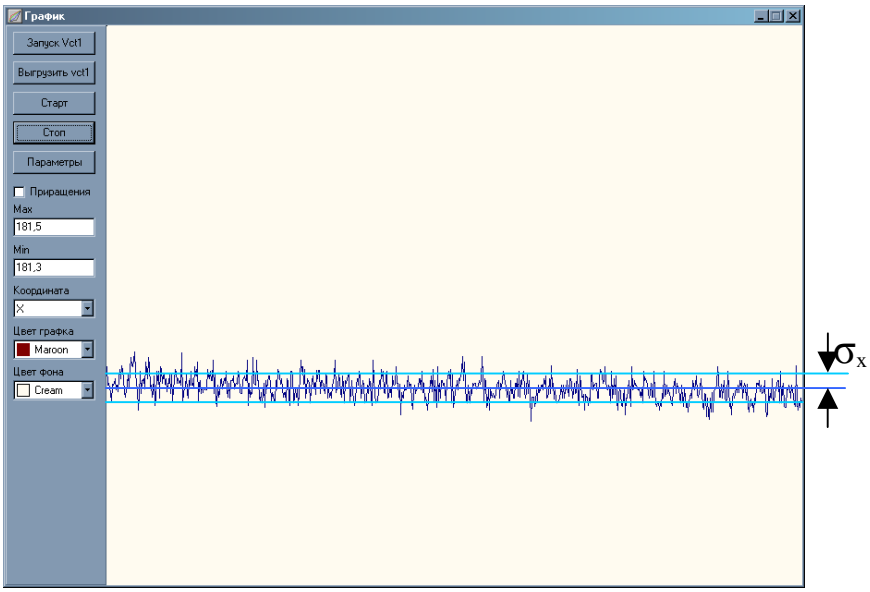


Fig. A1.1 Random component of measuring error for coordinate X (HTS)

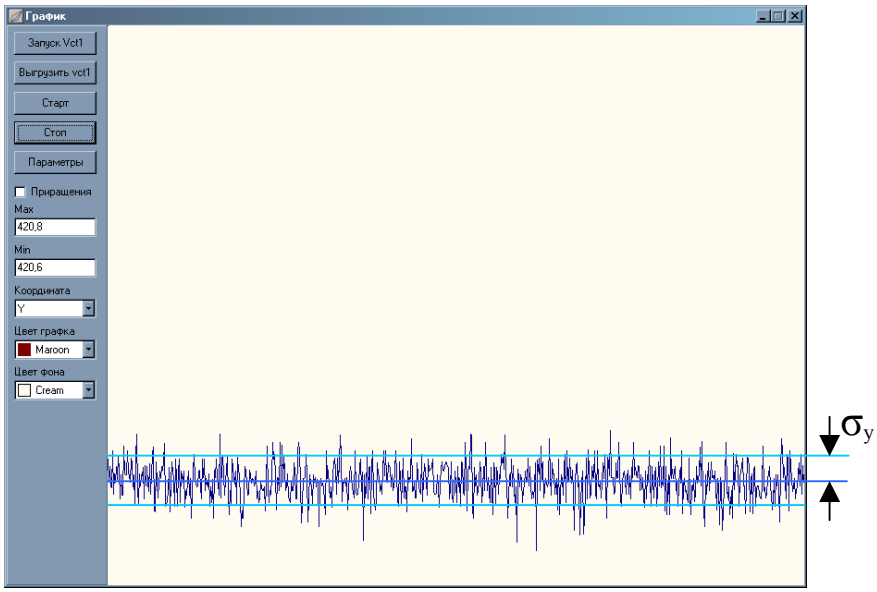


Fig. A1.2 Random component of measuring error for coordinate Y (HTS)

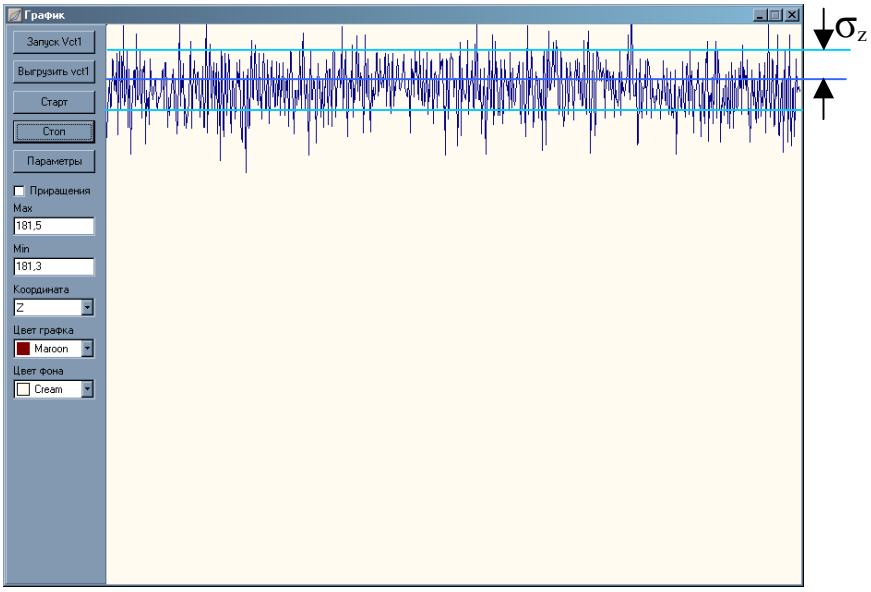


Fig. A1.3 Random component of measuring error for coordinate Z (HTS)

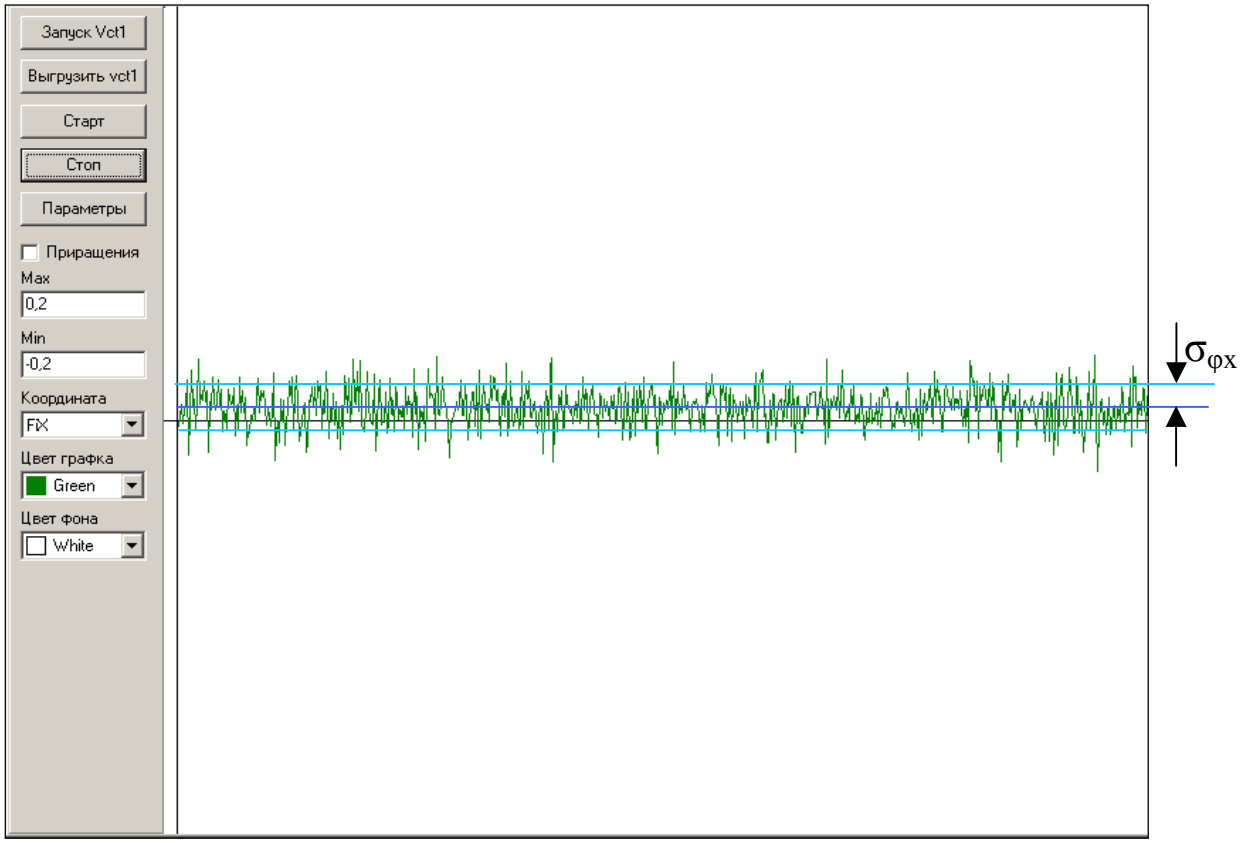


Fig. A1.4 New version of HTS prototype, RMS angle φ_x (yaw), full dagram size ± 0,2 ang.deg. (HTS)

The RMS error for HTS prototype after hardware and software adjastment is decreased more than 5 times as compared with previous version of HTS prototype (compare full size diagrams of RMS Fig. A1.4 and Fig. A1.5 ±0,2 ang.deg and ±1,0 ang.deg. respectively).

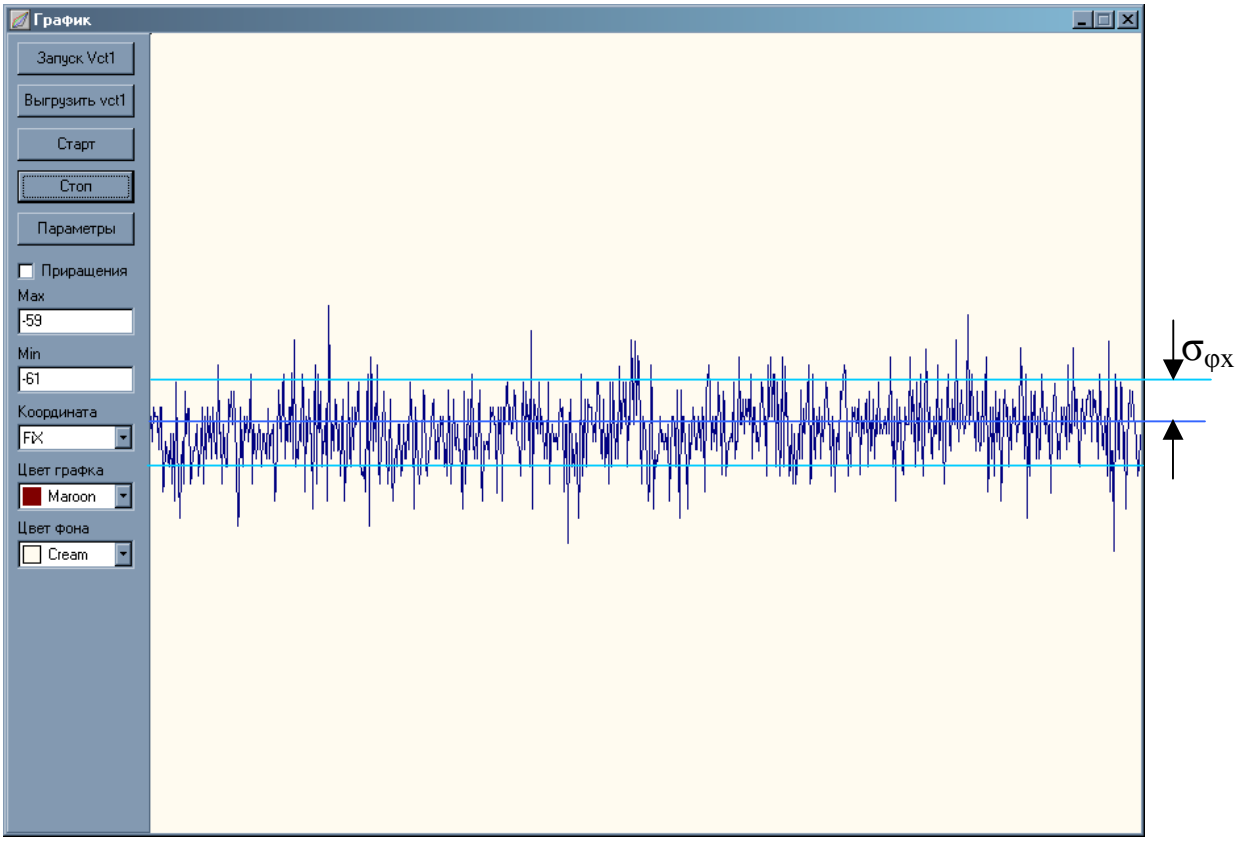


Fig. A1.5 Random component of measuring error for coordinate φ_x (previous version of HTS prototype)

2. Head motion box (HMB) measurement results for active HTS and HTS+

2.1. Head motion box (HMB) measurement results for active HTS

Table A1.3

it	Name of parameter	Units	X	Y	φ_x	φ_y	φ_z	Note
1.	Maximal zone at distance RDU to CU (Z=400 mm)	mm, deg.	250	240	±90	±40	±34	4-mark RDU
2.	Maximal zone at distance RDU to CU (Z=800 mm)	mm, deg.	660	500	±88	±33	±30	
3.	Maximal zone at distance RDU to CU (Z=1600 mm)	mm, deg.	1450	1050	±88	±30	±28	

2.2. Hand motion box (Hand MB) measurement results for active HTS+

Table A1.4

it	Name of parameter	Units	X	Y	φ_x	φ_y	φ_z	Note
1.	Maximal zone at distance RDU to CU (Z=200 mm)	mm, deg.	260	230	±90	±40	±35	3-mark RDU
2.	Maximal zone at distance RDU to CU (Z=800 mm)	mm, deg.	740	620	±85	±33	±30	
3.	Maximal zone at distance RDU to CU (Z=2000 mm)	mm, deg.	1600	1200	±80	±28	±30	

Work volume for hand control using HTS+ (Hand MB) is larger than HMB using HTS.
The experimental diagrams of RDU movements in three angles (yaw, pitch, roll) are shown below in Fig. A1.6 … A1.8 (below). The wide-angle range corresponds to the new design of RDU with 10 reference points.

3. Common results of coordinate measurement error (RMS) for HTS and HTS+

The calculated RMS values (σ) for HTS and HTS+ for 6D coordinates are presented in the table A1.5.

Random component of coordinate measurement error (RMS) for HTS and HTS +

Table A1.5

	HTS	HTS+
σ_x	0,008 pixels	0,009 pixels
σ_y	0,01 pixels	0,013 pixels
σ_z	0,02 mm	0,02 mm
$\sigma_{\varphi x}$	1,0 arc. min.	1,2 arc. min.
$\sigma_{\varphi y}$	2,2 arc. min.	2,8 arc. min.
$\sigma_{\varphi z}$	2,6 arc. min.	2,9 arc. min.

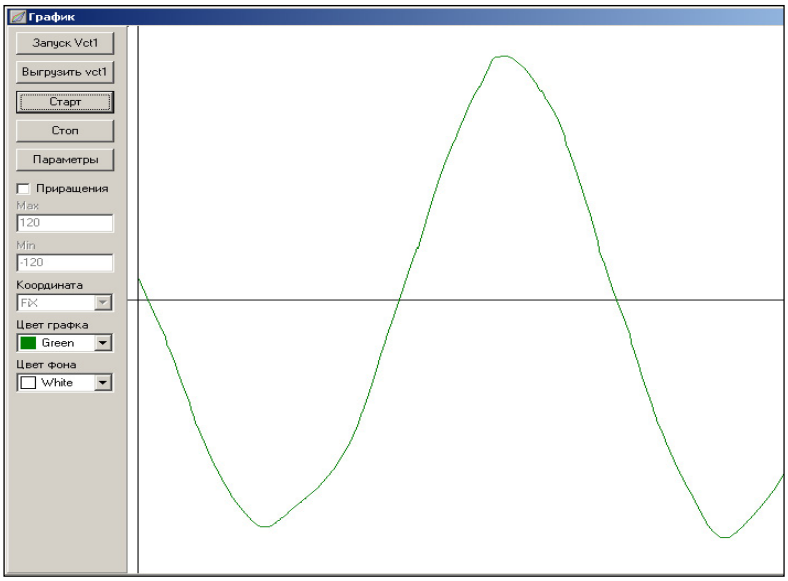


Fig. A1.6 HTS prototype, Range of angle φx (yaw), full diagram size ± 120 angle degree

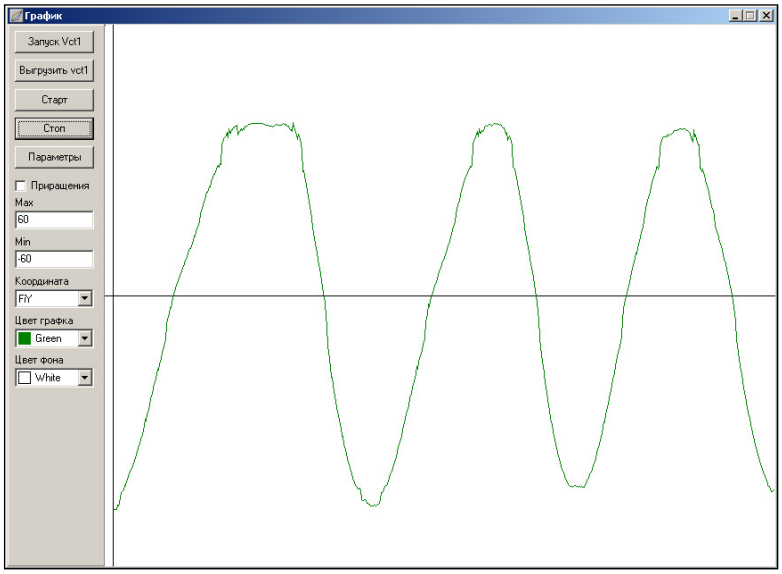


Fig. A1.7 HTS prototype, Range of angle φy (pitch), full diagram size ± 60 angle degree

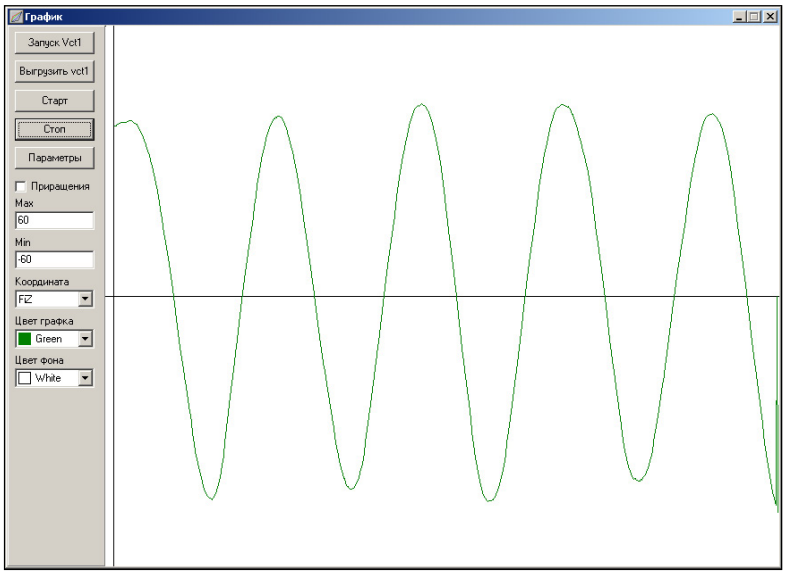


Fig. A1.8 HTS prototype, Range of angle φz (roll), full diagram size ± 60 angle degree

4. Response time in control mode for the Remote Robot-like device for Viewing environment (RRV) and Remote Robot-manipulator Control (RRC) using HTS and HTS+

4.1. Response time in control mode for RRV using HTS (response time in ms)

Table A1.6

Prototype	HTS properly	RRV Control by HTS
1. Active HTS without filtering output data	40	300
2. Active HTS with filtering output data	200	430
3. Active HTS with prediction of output data	80	350

The experiments were carried out at different modes of HTS output data processing using median filtering (m=10) and predictive Kalman filter (k=5).

4.2. Response time in control mode for RRC using HTS+(response time in ms)

Table A1.7

Prototype	HTS+ properly	RRC Control by HTS+
1. Active HTS+ without filtering output data	40	240
2. Active HTS+ with filtering output data	200	400
3. Active HTS+ with prediction of output data	80	320

The experiments were carried out at different modes of HTS+ output data processing using median filtering (m=10) and predictive Kalman filter (k=5).

4.3. Response time in control mode for RRC using HTS with novel CCD camera (100 Hz frame rate)

Table A1.8

Prototype	HTS properly	RRC Control by HTS
1. Active HTS without filtering output data	5	40
2. Active HTS with filtering output data	25	220
3. Active HTS with prediction of output data	10	120

The experiments were carried out at different modes of HTS output data processing using median filtering (m=10) and predictive Kalman filter (k=5).

5. The experimental studies of the process RDU reference point identification

In that stage of experiments some variants of Reference Device Unit (RDU) for the HTS prototype were studied:
- RDU-3 with 3 reference points (HTS+);
- RDU-4 with 4 reference points (HTS);
- new version of RDU-10 with 10 reference points for 180 angle deg. of head azimuth rotation.

The HTS SW prototype with 3D wire-frame models of RDU was used for experimental study of identification methods.

Some results of identification for RDU-3 and RDU-10 presented below in Fig. A1.9 and Fig. A1.10 correspondently.
Additionally the filtering of sunlight interference during the identification of RDU image had been studied. In Fig. A1.11 an image of the Sun in HTS cameras FOV is shown.

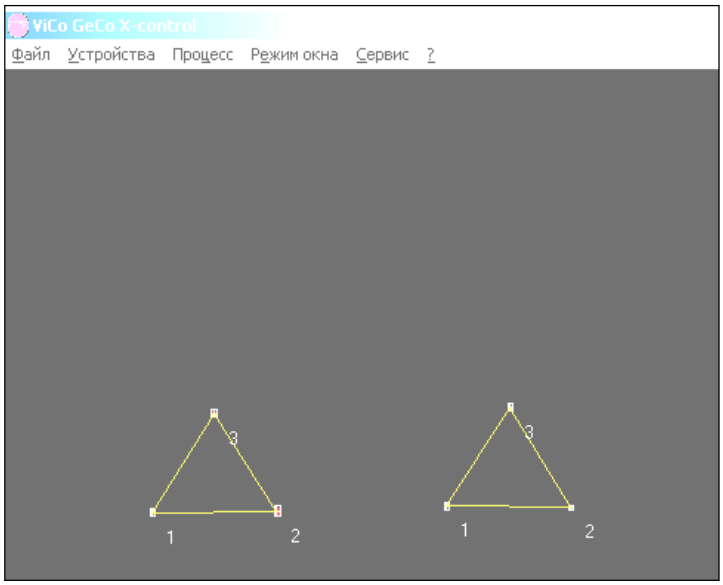


Fig. A1.9 The HTS prototype, identification of RDU-3 reference points (1-3).

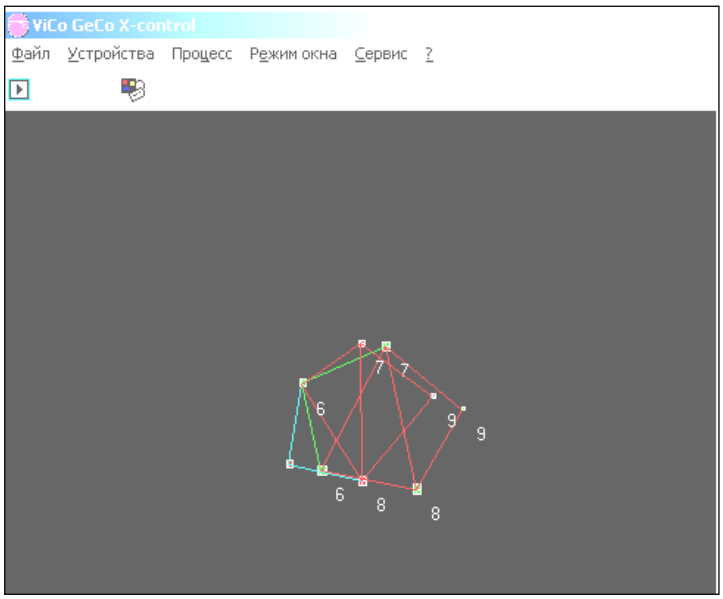


Fig. A1.10 The HTS prototype, identification of RDU-10 reference points (6-9).

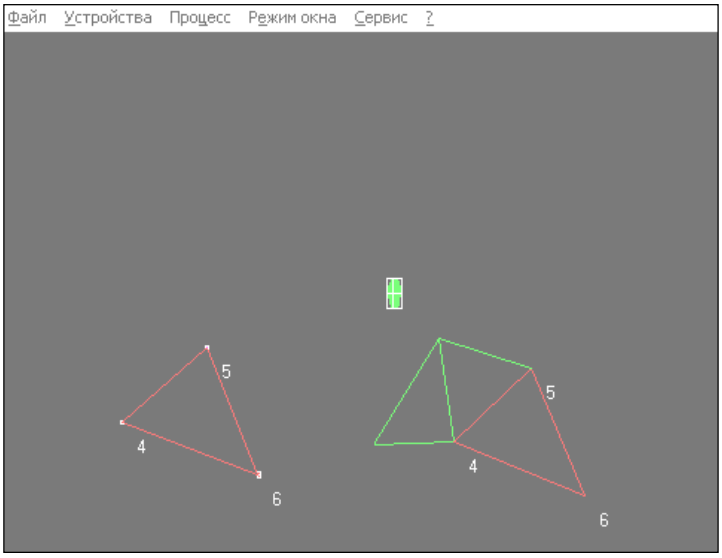


Fig. A1.11 HTS prototype, identification of RDU-10 reference points (4-6), with Sun light interference

6. The examples of 3D model descriptions

For examples, some RDU 3D models in tabular descriptions are presented below in Fig. A1.12 – Fig. A1.15.

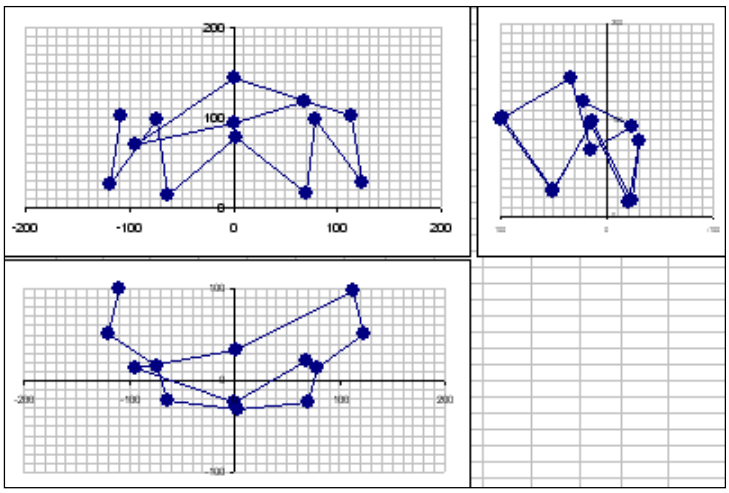
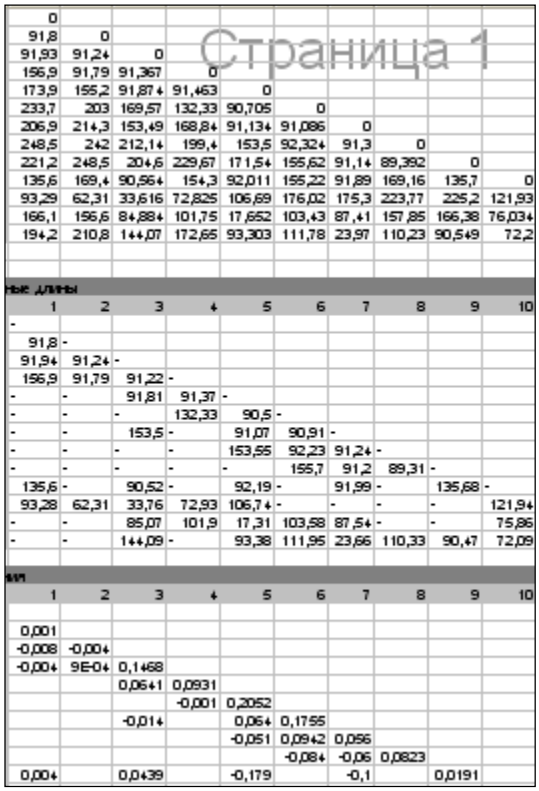


Fig. A1.12 The HTS prototype, 3D model RDU (10 reference points)



0									
91,8	0								
91,93	91,24	0							
156,9	91,79	91,367	0						
173,9	155,2	91,874	91,463	0					
233,7	203	169,57	132,33	90,705	0				
206,9	214,3	153,49	168,84	91,134	91,086	0			
248,5	242	212,14	199,4	153,5	92,324	91,3	0		
221,2	248,5	204,6	229,67	171,54	155,62	91,14	89,392	0	
135,6	169,4	90,564	154,3	92,011	155,22	91,89	169,16	135,7	0
93,29	62,31	33,616	72,825	106,69	176,02	175,3	223,77	225,2	121,93
166,1	156,6	84,884	101,75	17,652	103,43	87,41	157,85	166,38	76,034
194,2	210,8	144,07	172,65	93,303	111,78	23,97	110,23	90,549	72,2

	1	2	3	4	5	6	7	8	9	10
-										
	91,8	-								
	91,94	91,24	-							
	156,9	91,79	91,22	-						
-	-		91,81	91,37	-					
-	-		-	132,33	90,5	-				
-	-		153,5	-	91,07	90,91	-			
-	-		-	-	153,55	92,23	91,24	-		
-	-		-	-	-	155,7	91,2	89,31	-	
	135,6	-	90,52	-	92,19	-	91,99	-	135,68	-
	93,28	62,31	33,76	72,93	106,74	-	-	-	-	121,94
-	-		85,07	101,9	17,31	103,58	87,54	-	-	75,86
-	-		144,09	-	93,38	111,95	23,66	110,33	90,47	72,09

	1	2	3	4	5	6	7	8	9	10
	0,001									
	-0,008	-0,004								
	-0,004	9E-04	0,1468							
			0,0641	0,0931						
				-0,001	0,2052					
			-0,014		0,064	0,1755				
					-0,051	0,0942	0,056			
						-0,084	-0,06	0,0823		
	0,004		0,0439		-0,179		-0,1		0,0191	

J Left	I TK	Name Right	Type	Point1	Point2	Point3	b12	b13	b23	By	Bz	Bx	TK
1 No	1	123	LRT	2	3	1	61	91	90	-28	10	-82	Yes
1 No	2	234	VRT	2	4	3	91	60	71	57	-16,2	12,6	Yes
1 No	3	345	NRT	5	3	4	110	110	72	-200	-28	20	Yes
1 Yes	4	456	VRT	4	6	5	133	105	105	0	-9,4	3,4	Yes
1 No	5	567	NRT	7	5	6	25	115	95	28	180	170	-
1 No	6	678	VRT	6	8	7	90	110	110	-53	-21,8	-4,2	Yes
1 No	7	789	PRT	7	8	9	110	91	91	25	5	90	Yes
2 No	1	103	PRT	10	1	3	135,	120	90	0	0	0	-
2 No	3	305	VRT	3	5	10	110	120	90	0	0	0	Yes
2 No	5	507	VRT	5	7	10	91	91	92	0	0	0	-
2 No	7	709	LRT	7	9	10	90	90	136	0	0	0	Yes

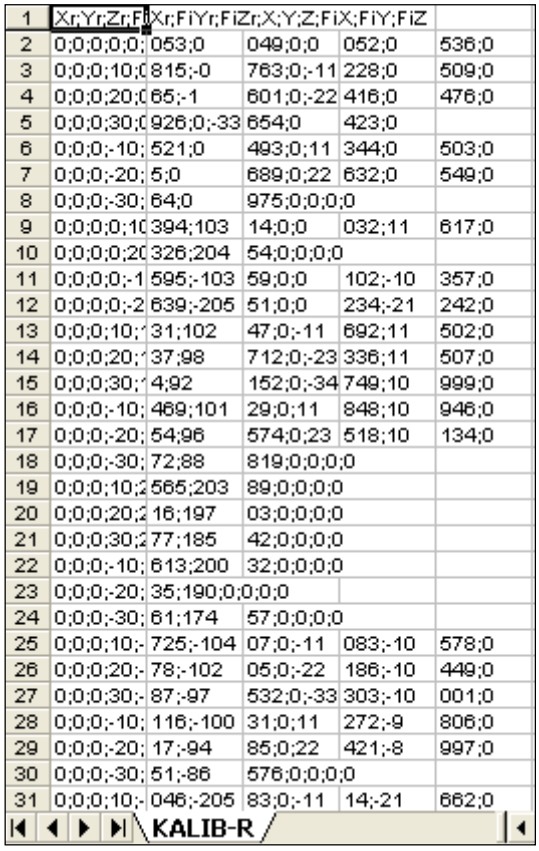
Fig. A1.14 The HTS prototype identification listing of RDU (10 reference points)

103		305		507		709	Name
0=1 3=3 1=2		0=3 3=1 5=2		0=3 5=1 7=2		0=2 7=3 9=1	Numeration of Points
135,5		92		91		136	b12 (mm)
90,5		90,5		91		91	B13
92,5		91		92		92	B23
- 166		32		- 32		166	My (yaw)
- 10		18 40’		18 40’		-10	Mz (pitch)
30 40’		0		0		- 30 40’	Mx (roll)
123	234	345	456	567	678	789	Name
3=2 1=3 2=1	3=3 2=1 4=2	3=2 5=1 4=3	5=3 4=1 6=2	5=2 7=1 6=3	7=3 6=1 8=2	7=1 9=3 8=2	Numeration of Points
91	91	92	132,5	91	91,5	90,5	b12 (mm)
91,5	91	91,5	91,5	91	91	91	B13
92,5	92	92	91,5	91,5	90,5	90,5	B23
- 10	50	- 58	0	58	- 50	10	My (yaw)
10	58 40’	- 54 40’	56 20’	- 54 40’	58 40’	10	Mz (pitch)
-61	- 13	180	0	180	13	61	Mx (roll)

Fig. A1.15 The HTS prototype identification table of RDU (10 reference points)

7. Examples of calibration data for HTS prototype

As an illustration of the experimental studies of calibration process for HTS prototype’s cameras data fragments are shown in Fig. A1.16 and A1.17.



1	Xr;Yr;Zr;Fi	Xr;FiYr;FiZr;X;Y;Z;FiX;FiY;FiZ			
2	0;0;0;0;0;	053;0	049;0;0	052;0	536;0
3	0;0;0;10;0	815;-0	763;0;-11	228;0	509;0
4	0;0;0;20;0	65;-1	601;0;-22	416;0	476;0
5	0;0;0;30;0	926;0;-33	654;0	423;0	
6	0;0;0;-10;	521;0	493;0;11	344;0	503;0
7	0;0;0;-20;	5;0	689;0;22	632;0	549;0
8	0;0;0;-30;	64;0	975;0;0;0;0		
9	0;0;0;0;10	394;103	14;0;0	032;11	617;0
10	0;0;0;0;20	326;204	54;0;0;0;0		
11	0;0;0;0;-1	595;-103	59;0;0	102;-10	357;0
12	0;0;0;0;-2	639;-205	51;0;0	234;-21	242;0
13	0;0;0;10;1	31;102	47;0;-11	692;11	502;0
14	0;0;0;20;1	37;98	712;0;-23	336;11	507;0
15	0;0;0;30;1	4;92	152;0;-34	749;10	999;0
16	0;0;0;-10;	469;101	29;0;11	848;10	946;0
17	0;0;0;-20;	54;96	574;0;23	518;10	134;0
18	0;0;0;-30;	72;88	819;0;0;0;0		
19	0;0;0;10;2	565;203	89;0;0;0;0		
20	0;0;0;20;2	16;197	03;0;0;0;0		
21	0;0;0;30;2	77;185	42;0;0;0;0		
22	0;0;0;-10;	613;200	32;0;0;0;0		
23	0;0;0;-20;	35;190;0;0;0;0			
24	0;0;0;-30;	61;174	57;0;0;0;0		
25	0;0;0;10;-	725;-104	07;0;-11	083;-10	578;0
26	0;0;0;20;-	78;-102	05;0;-22	186;-10	449;0
27	0;0;0;30;-	87;-97	532;0;-33	303;-10	001;0
28	0;0;0;-10;	116;-100	31;0;11	272;-9	806;0
29	0;0;0;-20;	17;-94	85;0;22	421;-8	997;0
30	0;0;0;-30;	51;-86	576;0;0;0;0		
31	0;0;0;10;-	046;-205	83;0;-11	14;-21	662;0

KALIB-R

Fig. A1.16 The HTS prototype a fragment of calibration table of cameras

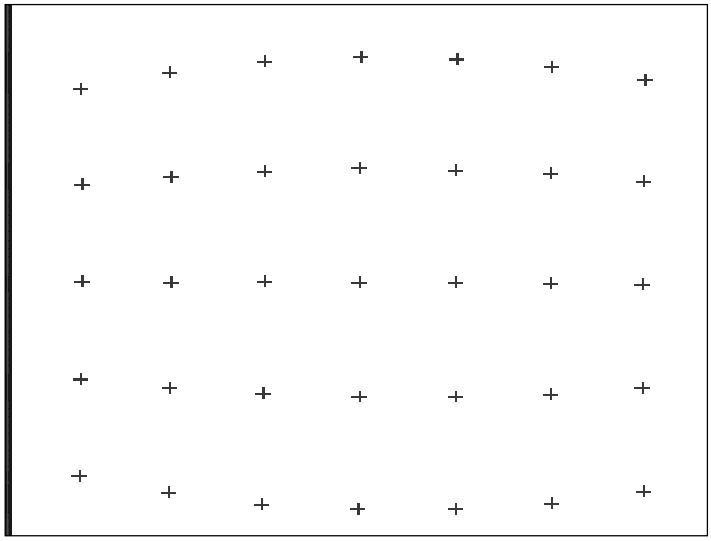


Fig. A1.17 The HTS prototype distortions table of cameras CCU

8. Experiments of teaching by show with HTS+

The experiments were carried out using 3-mark reference device HTS +for teaching by show of robot-manipulator. The some results of experiments with teaching of 3D-trajectory (Xr, Yr, Zr) and storing in robot memory are presented below.

The memorized data of HTS+ are presented (with filtering) on Table A1.9 and Fig. A1.18.

Table A1.9

Xr	Yr	Zr
2,0	0,1	287,9
8,0	66,1	287,6
4,0	131,3	286,9
-2,0	194,6	285,4
0,1	-65,6	287,4
0,9	-129,4	286,6
5,0	-190,5	284,8
10,2	0,1	389,5
20,5	0,3	488,4
-10,7	0,4	185,7
-20,9	0,3	83,7
-10,3	67,7	186,1
-10,1	132,8	188,8
-10,5	196,5	193,7
-10,5	-65,9	187,2
-10,8	-130,3	190,9
-10,1	-192	196,9

Xr	Yr	Zr
-20,2	133,5	91,3
-20,4	197,4	102,6
-20,7	-66,1	87,1
-20,9	-130,9	95,6
-20,8	-192,9	109,1
10,5	66,3	388,7
10,4	131,7	384,7
10,7	195	377
10,9	-65,6	387,3
10,8	-129,5	381,5
10,6	-190,8	372,9
20,4	66,5	486,7
20,9	132,3	479,0
20,4	195,8	465,5
20,7	-65,6	483,6
20,1	-129,9	473,7
20,0	-191,4	457,9

The 3D movement teaching before execution with robot-manipulator must be compressed and filtered. Additional processing of teaching trajectory with usage of recognition methods and 3D frame-structural model (FSM) has been studied in this Project Task 6 too [26].

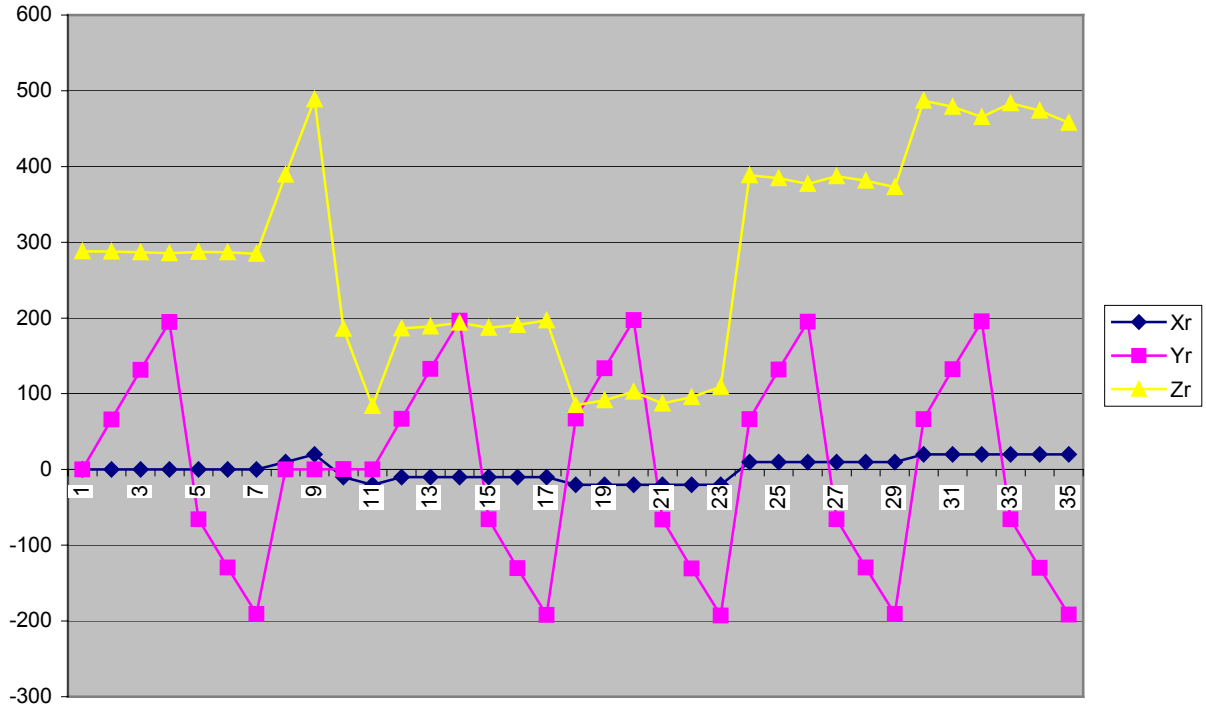


Fig. A1.18 The HTS+ prototype diagram of teaching by show

The preliminary Results of identification procedure for 3D teaching movement are presented on Fig. A1.19. The primitives of shape movement (motion pattern) are recognized: Segment (S), Step Up (SU), Step Down (SD), Peak (P), Zigzag (Z), Arc (A), Hollow (H).

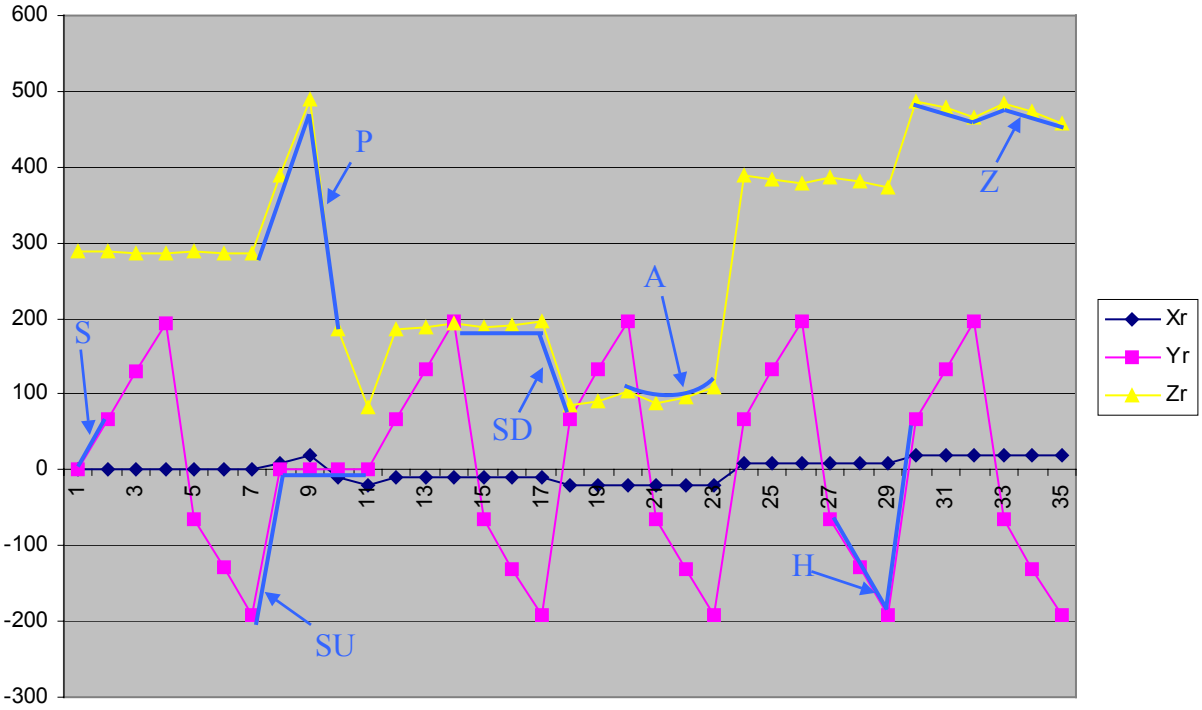


Fig. A1.19 The recognized primitives of shape movement (motion pattern)

S - Segment, SU - Step Up, SD - Step Down, P - Peak, Z - Zigzag, A- Arc, H - Hollow

Appendix 2

Results of experimental testing the ETS prototype

Some results of experimental testing the ETS prototype are presented in this Appendix. The ETS data are presented on diagrams Fig. A2.1-A2.4 below.

1. Estimate of ETS RMS error

The results of experiments with ETS prototype, RMS error see in Table A2.1.

Table A2.1

it	Name of parameter	Units	φ_x	φ_y	Note
1.	RMS (σ) error (without filtering)	Arc. min	8	12	ETS-R (right eye)
2.	RMS (σ) error (with median filtering)	Arc. min	6	10	
3.	RMS (σ) error (without filtering)	Arc. min	6	10	ETS-L (left eye)
4.	RMS (σ) error (with median filtering)	Arc. min	7	8	

2. The range of angles for ETS prototype

The results of experiments with ETS prototype for angles of eye rotation range in horizontal and vertical directions, see in Table A2.2.

Table A2.2

it	Name of parameter	Units	φ_x	φ_y	Note
1.	Maximal range of right eye rotation, measured by ETS	deg.	±22	±18	ETS-R (right eye)
2.	Maximal range of left eye rotation, measured by ETS	deg.	±20	±15	ETS-L (left eye)

3. The ETS prototype output data time delay

The results of experiments with ETS prototype, time delay of ETS output data in horizontal and vertical directions, see in Table A2.3.

Table A2.3

it	Name of parameter		Units	φ_x	φ_y	Note
1.	Time delay of ETS prototype output data	Right eye	ms	40	40	ETS-R (right eye)
2.	Time delay of ETS prototype output data with median filtering (m=5)		ms	120	120	ETS-R (right eye)
3.	Time delay of ETS prototype output data with predictive Kalman filtering		ms	50	50	ETS-R (right eye)
4.	Time delay of ETS prototype output data	Left eye	ms	40	40	ETS-L (left eye)
5.	Time delay of ETS prototype output data with median filtering (m=5)		ms	120	120	ETS-L (left eye)
6.	Time delay of ETS prototype output data with predictive Kalman filtering		ms	50	50	ETS-L (left eye)

The experiments were carried out at different modes of ETS output data processing using median filtering (m=5) and predictive Kalman filter (k=5).

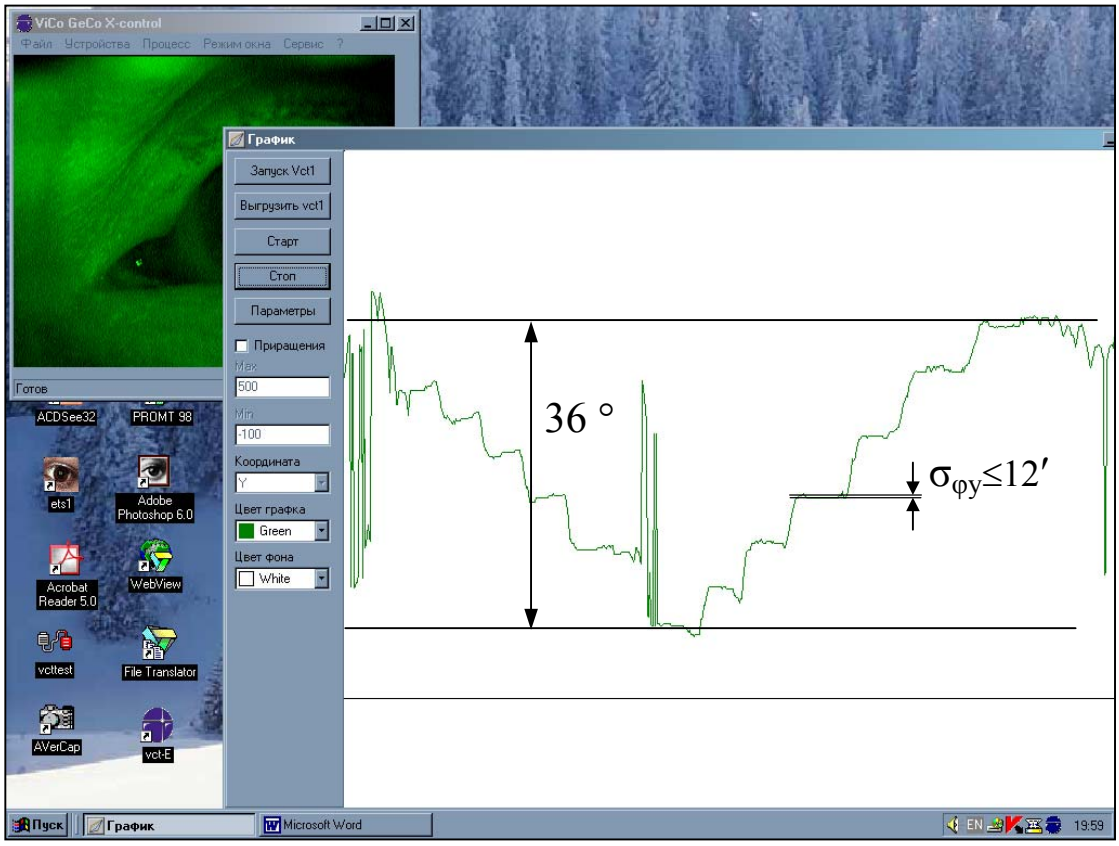


Fig. A2.1 ETS output data of eye turns at discrete angles in vertical direction (coordinate φ_Y)

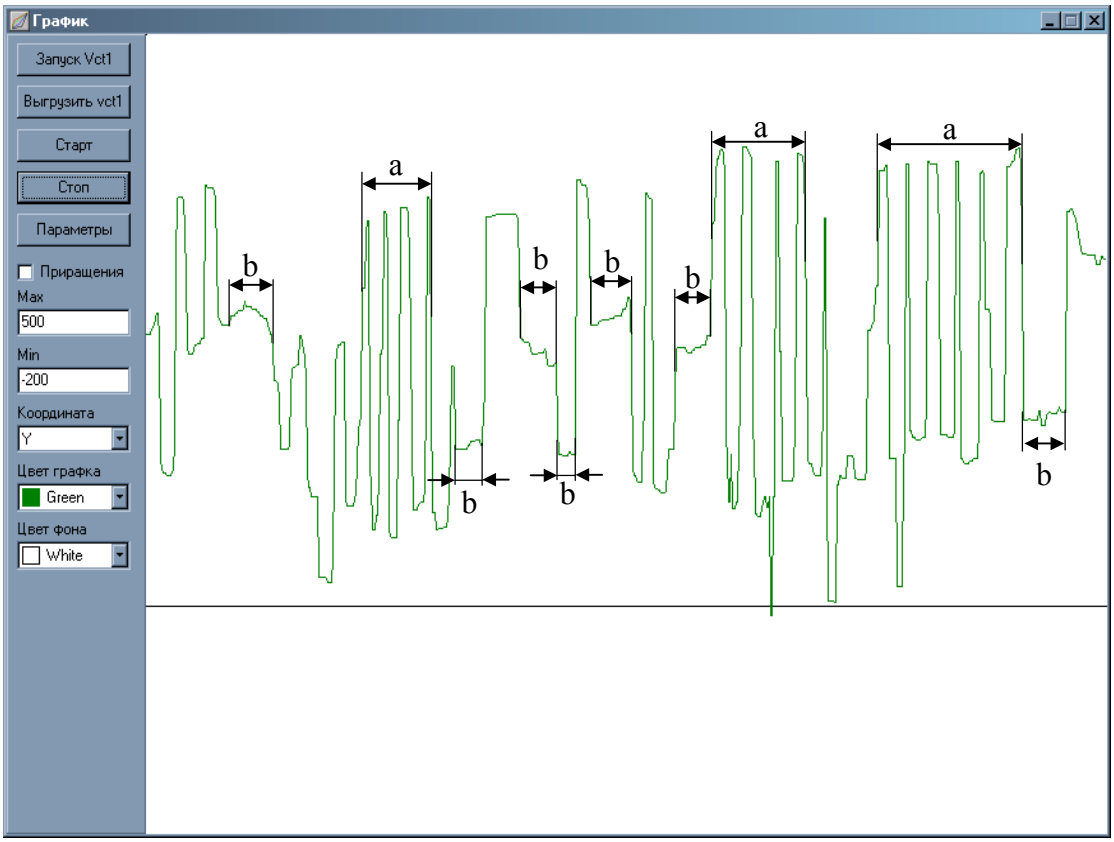


Fig. A2.2 ETS output data range for eye turn in vertical direction (coordinate φ_Y)

a – Time intervals of eye saccades from one extreme position to another;

b - Time interval of look fixing.

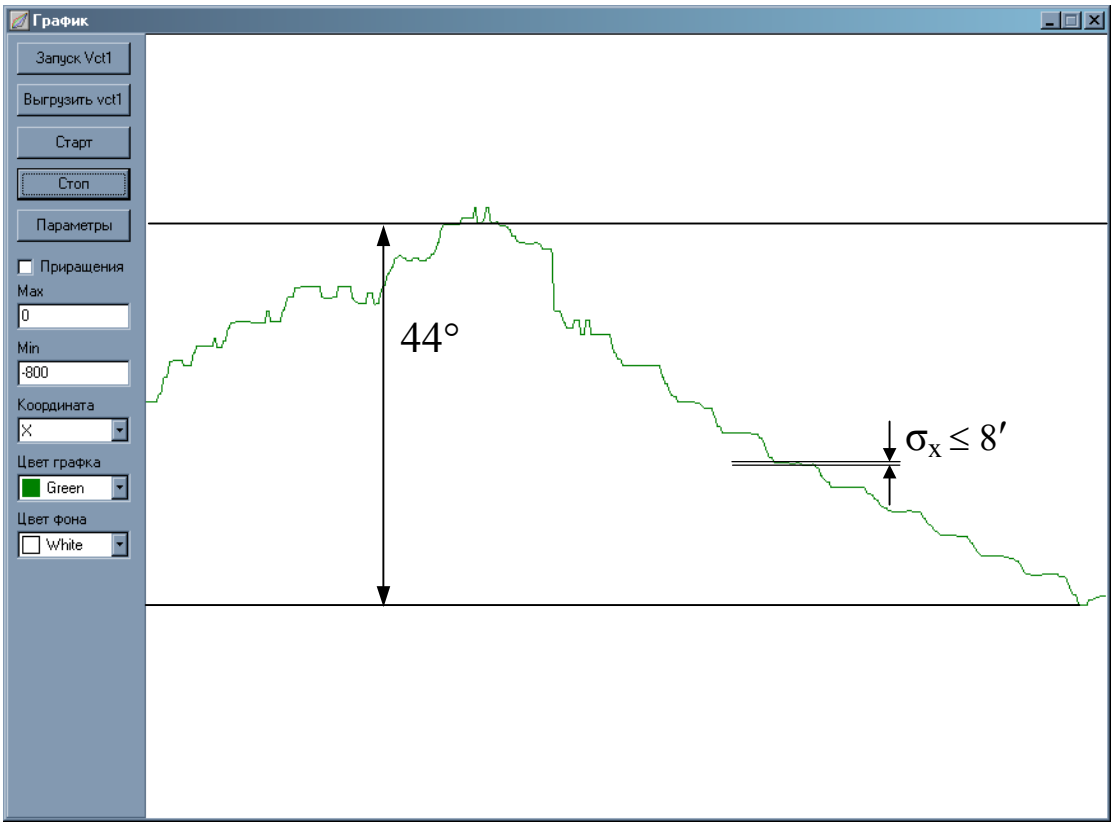


Fig. A2.3 ETS output data of eye turns at discrete angles in horizontal direction (coordinate φ_X)

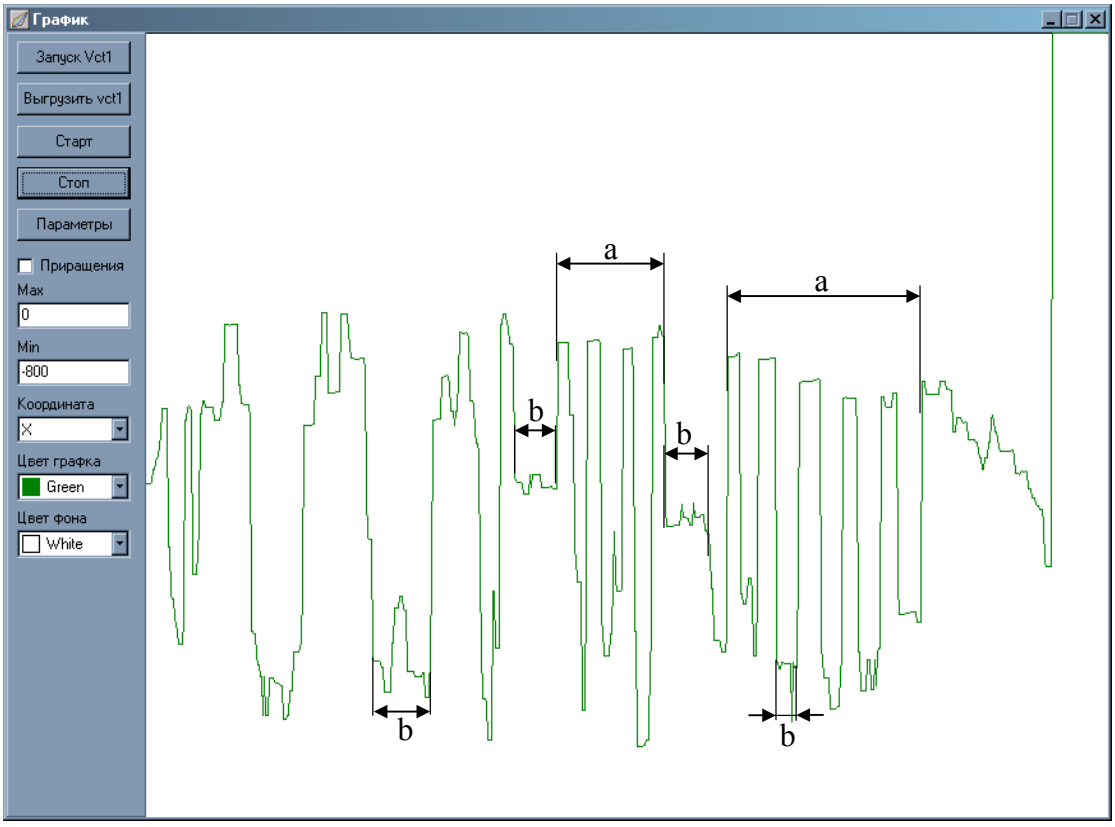


Fig. A2.4 ETS output data range for eye turn in horizontal direction (coordinate φ_X)
a – Time intervals of eye saccades from one extreme position to another;
b - Time interval of look fixing.

4. The experiments with HRIZ controlled ETS for video frames

Some results of experimental testing the ETS prototype for HRIZ control of video frames a presented on Fig. A2.5 (video of 100 frames).

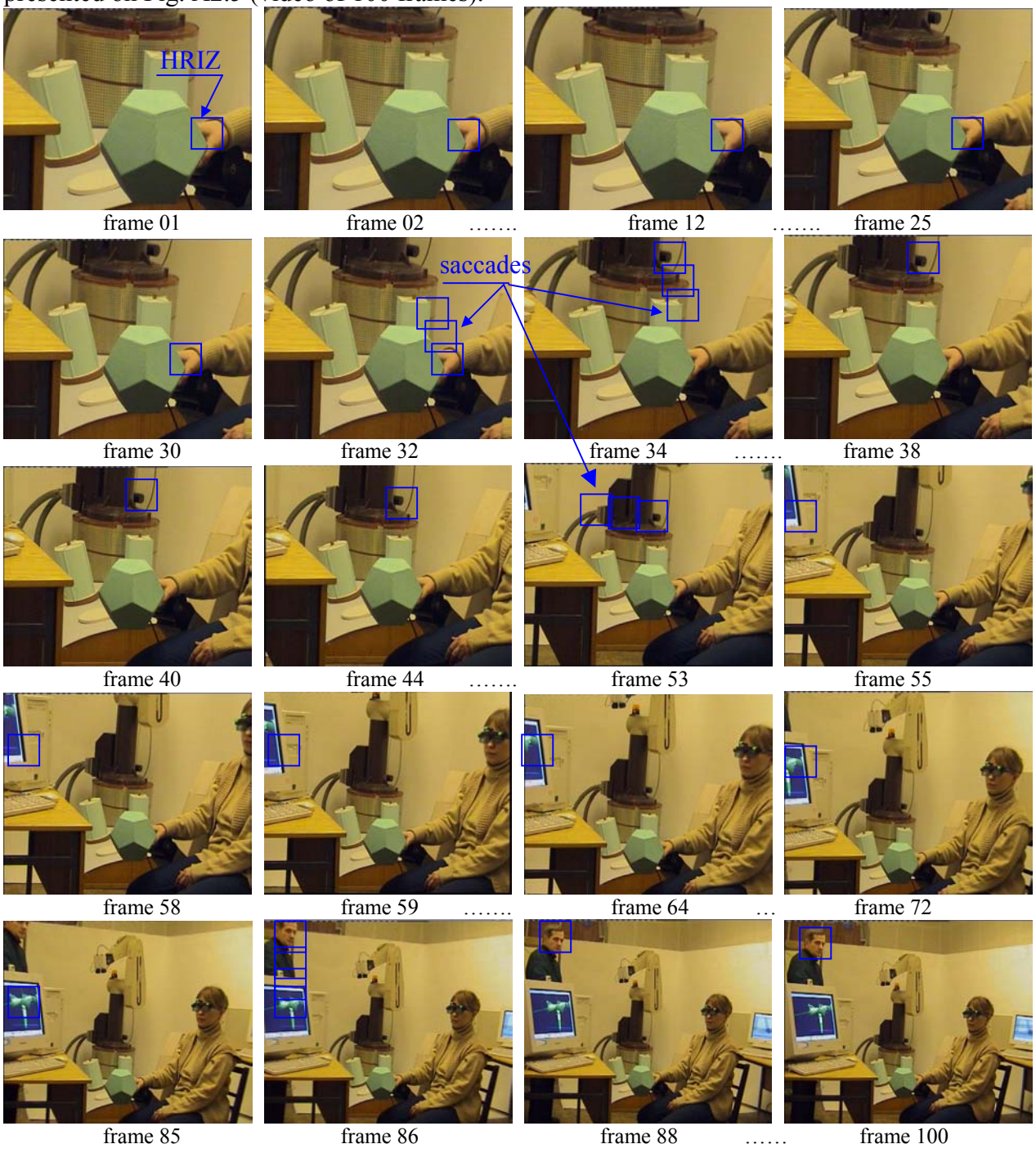


Fig. A2.5 Testing the ETS prototype for HRIZ control of video frames

The preliminary test results of main parameters ETS prototype for HRIZ control of video:

1. The estimating drift of gaze position – 1-3% of picture size.
2. The size of HRIZ for distance 0,5 meter to display (19") – 10-15% of picture size.
3. The resolution in HRIZ – 1240 x 1024.
4. The resolution in peripheral zone 320 x 240.
5. Middle time interval of the gaze fixation – 0,5 – 5 sec (10 –100 frames).
6. Middle time interval of the gaze saccade – 0,05 – 0,2 sec (2 – 5 frames).

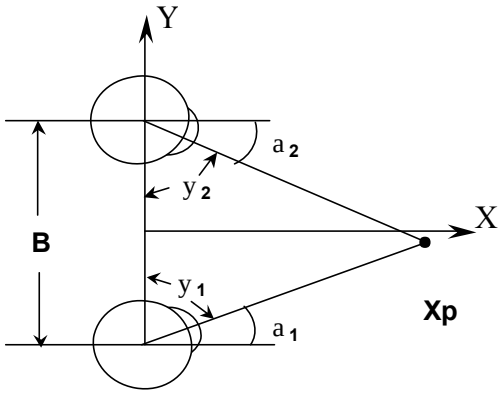
The compression for video with ETS control of HRIZ was estimated as 8-16.

Appendix 3

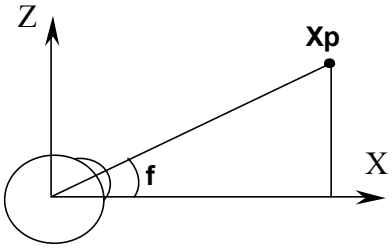
Biotechnical Eye Scanning

In this appendix, we describe an original method of obtaining input data for the generation of a geometrical model of the robot environment. This is biotechnical eye scanning, in which input data are generated as a 3D point array on the surface separating the transparent and opaque regions of the robot working zone. The operator visually scans by eyes, in a very natural way, either the working zone itself or its TV image.

The possibility to generate a 3D point array is based on the mechanisms underlying the operation of the vision region in a human brain. When eyes are focused at an object being examined or at its fragment, the regulatory signals going from the brain to the muscles responsible for the motion of each eye ball make them contract in a specific way. The contractions change the angular values of the left α_1, φ_1 and right α_2, φ_2 eye ball, so that the optical eye axes intercept at the point 'of interest' in a surface region of the fragment being examined (this point lies at the region center, see Fig. A3.1).



a)



b)

Fig. A3.1 Eyes focused at an object X_p
a). vertical plane view b). side view

Clearly, the axial sections limited by the interception point at one end and by the eye ball rotation centers at the other end (IPD) form a triangle (Fig. A3.1).

Two angles of the triangle, ψ_1 and ψ_2, are the linear functions of the values α_1 and α_2 of the eyeball angular displacements. If one knows the values of α_1, φ_1 and α_2, φ_2, measured by EHS, the IPD value B, the position of the point of interest in the head axes coordinate system can be found as

$$x_1 = \frac{B}{tg\alpha_2 - tg\alpha_1}; \qquad x_2 = \frac{B}{2} + x_1 tg\alpha_1; \qquad x_3 = x_1 \frac{tg\varphi_1}{\cos\alpha_1}$$

In order to find the position of this point in the inertial coordinate system, it is sufficient to know the head position and orientation in the same coordinates.

Therefore, the generation of a 3D point array on the surface separating the opaque and transparent regions of the working zone visualized in real time requires the use of the following hardware.

If the operator visualizes the working zone directly, the hardware is to include

(1) an eye tracking system, whose possible variant was described in Interim Report I, Task 5; during the visualization of a working zone this system generates the values of the eyeball angular displacements with the updating period of 40 ms;
(2) a head tracking system, whose possible modifications were described in Interim Report, Task 5; during the visualization this system generates linear and angular coordinates determining the head position/orientation in the working zone axial coordinates; the updating period is about 10 ms, a much smaller value than in the eye tracking system;
(3) a computer to process eye and head tracking data as well as to calculate the coordinates of 3D points on the geometrical model surface in the inertial coordinates;

During the eye scanning of a working zone with a periodicity of about 40 ms, this hardware provides sets of 10 values including the four angles α_1, α_2, φ_1 and φ_2, which determine the direction of each eye, and six coordinates, which determine the head position/orientation in the inertial coordinate frame. These sets of values are used to find the sequence of the points of interest for the generation of the environment geometrical model in the inertial coordinate system. The operator is continuously involved in the generation of such point arrays by scanning the working zone visually, and this explains the term biotechnical eye scanning.

If the operator visualizes the working zone with TV cameras, the hardware must include, in addition to the equipment listed above,

(1) two coupled TV cameras with the IPD equal to the human IPD; the cameras are placed at the sites of the operator's eyes and their function is to produce TV-stereo images of the working zone;
(2) a helmet with two mounted displays: one for getting video images from the left TV camera and the other from the right TV camera.

By making a proper calibration, one can get the necessary TV-image scale. The scale must be such that the angular eye displacements during the TV-scanning of the zone fragments are exactly equal to the angles necessary for the scanning of the same fragments with a naked eye. If this condition is met, the operator can scan stereo-images of the working zone instead of scanning the real zone. In contrast to the case above, one can generate a cursor on each of the helmet displays. The cursor positions (x, y) on the displays must be defined as

$$x_i = l \sin \alpha_i; \quad y_i = l \sin \beta_i$$

where l is the distance between the eye ball center and the screen and i is the display number. The cursor will mark the point of interest on the left display for the left eye and on the right display for the right eye. Clearly, the operator will see a stereo-image of the cursor which marks the working zone fragment being examined by the operator. This makes the eye scanning more meaningful.

The implementation of this idea will primarily depend on how soon the required measurement accuracy can be achieved. Nevertheless, the fact that a large point array can be measured within an acceptable time period and that obvious measurement errors due to the specificity of eye scanning can be avoided makes one to hope for getting encouraging results.

It is important if the operator sees the working zone with TV cameras. The point "of interest" position may be defined as function of the cursor position and of binocular parallax. This fact eliminates point "of interest" position error generated by the measurement errors of α_1, α_2, φ_1, φ_2 angles.